Estimation of Tail Risk based on Extreme Expectiles

Abdelaati Daouia[a], Stéphane Girard[b] and Gilles Stupfler[c,d]

[a] Toulouse School of Economics, University of Toulouse Capitole, France
[b] INRIA Grenoble Rhône-Alpes / LJK Laboratoire Jean Kuntzmann - MISTIS
[c] Aix Marseille Université, CNRS, EHESS, Centrale Marseille, GREQAM UMR 7316, 13002 Marseille, France
[d] School of Mathematical Sciences, The University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

Abstract

We use tail expectiles to estimate alternative measures to the Value at Risk (VaR), Expected Shortfall (ES) and Marginal Expected Shortfall (MES), three instruments of risk protection of utmost importance in actuarial science and statistical finance. The concept of expectiles is a least squares analogue of quantiles. Both expectiles and quantiles were embedded in the more general class of M-quantiles as the minimizers of an asymmetric convex loss function. It has been proved very recently that the only M-quantiles that are coherent risk measures are the expectiles. Moreover, expectiles define the only coherent risk measure that is also elicitable. The elicitability corresponds to the existence of a natural backtesting methodology. The estimation of expectiles did not, however, receive yet any attention from the perspective of extreme values. The first estimation method that we propose enables the usage of advanced high quantile and tail index estimators. The second method joins together the least asymmetrically weighted squares estimation with the tail restrictions of extreme-value theory. A main tool is to first estimate the large expectile-based VaR, ES and MES when they are covered by the range of the data, and then extrapolate these estimates to the very far tails. We establish the limit distributions of the proposed estimators when they are located in the range of the data or near and even beyond the maximum observed loss. We show through a detailed simulation study the good performance of the procedures, and also present concrete applications to medical insurance data and three large US investment banks.

Keywords: Asymmetric squared loss; Coherency; Expected shortfall; Expectiles; Extrapolation; Extreme values; Heavy tails; Marginal expected shortfall; Value at Risk.

1 Introduction

The concept of expectiles is a least squares analogue of quantiles, which summarizes the underlying distribution of an asset return or a loss variable Y in much the same way that quantiles do. It is a natural generalization of the usual mean $\mathbb{E}(Y)$, which bears the same relationship to this noncentral moment as the class of quantiles bears to the median. Both expectiles and quantiles are found to be useful descriptors of the higher and lower regions of the data points in the same way as the mean and median are related to their central behavior. Koenker and Bassett (1978) elaborated an absolute error loss minimization framework to define quantiles, which successfully extends the conventional definition of quantiles as left-continuous inverse functions. Later, Newey and Powell (1987) substituted the "absolute deviations" in the asymmetric loss function of Koenker and Bassett with "squared deviations" to obtain the population expectile of order $\tau \in (0,1)$ as the minimizer

$$\xi_\tau = \mathrm{argmin}_{\theta\in\mathbb{R}} \mathbb{E}\left\{\eta_\tau(Y-\theta) - \eta_\tau(Y)\right\}, \tag{1}$$

where $\eta_\tau(y) = |\tau - \mathbb{1}(y \leqslant 0)|\, y^2$, with $\mathbb{1}(\cdot)$ being the indicator function. Although formulated using a quadratic loss, problem (1) is well-defined as soon as $\mathbb{E}|Y|$ is finite, thanks to the presence of the term $\eta_\tau(Y)$. The first advantage of this asymmetric least squares approach relative to quantiles lies in the computational expedience of sample expectiles using only scoring or iteratively-reweighted least squares (see the R package 'expectreg'). The second advantage following Newey and Powell (1987) and Sobotka and Kneib (2012), among others, is that expectiles make more efficient use of the available data since the weighted least squares rely on the distance to data points, while empirical quantiles only utilize the information on whether an observation is below or above the predictor. Furthermore, sample expectiles provide a class of smooth curves as functions of the level τ, which is not the case for sample quantiles (see, *e.g.*, Schulze Waltrup *et al.* (2015)). Perhaps most importantly, inference on expectiles is much easier than inference on quantiles as already established by Newey and Powell (1987) and Abdous and Remillard (1995).

Value at Risk (VaR), Expected Shortfall (ES) and Marginal Expected Shortfall (MES) are three instruments of risk protection of utmost importance in actuarial science and statistical finance. They are traditionally based on the use of tail quantiles as a main tool for quantifying the riskiness implied by the great variability of losses and the heavy tails of their distribution. In this article we focus on the less-discussed problem of estimating the concepts of VaR, ES and MES when quantiles are replaced therein by expectiles. The use of expectiles as an alternative tool for quantifying tail risk has recently attracted a lot of interest, see for instance Martin (2014). A first motivating advantage of expectiles, following Kuan *et al.* (2009), is that they are more alert than quantiles to the magnitude of infrequent catastrophic losses. Also, they depend on both the tail realizations of Y and their probability, while quantiles only depend on the frequency of tail realizations and not on their values (Kuan *et al.* (2009)). Both families of quantiles and expectiles were embedded in the more general class of M-quantiles defined by Breckling and Chambers (1988) as the minimizers of an asymmetric convex loss function. Bellini (2012) has shown that expectiles with $\tau \geqslant \frac{1}{2}$ are the only M-quantiles that are isotonic with respect to the increasing convex order. Most importantly from the point of view of the axiomatic theory of risk measures, Bellini *et al.* (2014) have proved that the only M-quantiles that are coherent risk measures are the expectiles. They have also established that expectiles are robust in the sense of lipschitzianity with respect to the Wasserstein metric. Very recently, Ziegel (2016) has proved that expectiles are the only coherent law-invariant measure of risk which is also elicitable. The property of elicitability corresponds to the existence of a natural backtesting methodology. It has been shown that ES, the most popular coherent risk measure, is not elicitable (Gneiting, 2011), but jointly elicitable with VaR (Fissler and Ziegel, 2016).

In terms of interpretability, the τ-quantile determines the point below which $100\tau\%$ of the mass of Y lies, while the τ-expectile specifies the position ξ_τ such that the average distance from the data below ξ_τ to ξ_τ itself is $100\tau\%$ of the average distance between ξ_τ and all the data, *i.e.*,

$$\tau = \mathbb{E}\left\{|Y - \xi_\tau| \mathbb{1}(Y \leqslant \xi_\tau)\right\} / \mathbb{E}\left|Y - \xi_\tau\right|. \tag{2}$$

Thus, as pointed out by Fan and Gijbels (1996, p.231), the τ-expectile shares an intuitive interpretation similar to the τ-quantile, replacing the number of observations by the distance. This reduced 'probabilistic' interpretability of expectiles should not be viewed as a serious disadvantage however, since Bellini and Di Bernardino (2015) provide a transparent financial meaning of expectiles in terms of their acceptance sets: the τ-expectile defines the amount of money that should be added to a position in order to have a prespecified, sufficiently high gain-loss ratio. The gain-loss ratio is a popular performance measure in portfolio management and is well-known in the literature on no good deal valuation in incomplete markets (see Bellini and Di Bernardino (2015) and the references therein). Also, Ehm *et al.* (2016) have shown that expectiles are optimal decision thresholds in binary investment problems with fixed cost basis and differential taxation of profits versus losses. Another potential advantage for the adoption of expectiles in risk management, according to Taylor (2008), is that they are very closely related to the classical mean and the popular ES. Furthermore, the theoretical and numerical results obtained by Bellini and Di Bernardino (2015) seem to indicate that expectiles are perfectly reasonable alternatives to standard quantile-based VaR and ES. The statistical problem of expectile estimation did not, however, receive yet any attention from the perspective of extreme values.

Although least asymmetrically weighted squares estimation of expectiles dates back to Newey and Powell (1987) in case of linear regression, it recently regained growing interest in the context of nonparametric, semiparametric and more complex models, see for example Sobotka and Kneib (2012) and the references therein, as well as the two recent contributions by Holzmann and Klar (2016) and Krätschmer and Zähle (2016) for advanced theoretical developments. Attention has been, however, restricted to ordinary expectiles of fixed order τ staying away from the tails of the underlying distribution: in the latter two references, several asymptotic results such as uniform consistency and a uniform central limit theorem are shown for expectile estimators, but the order τ therein is assumed to lie within a fixed interval bounded away from 0 and 1. The purpose of this paper is to extend their estimation and asymptotic theory far enough into the tails. This translates into considering the expectile level $\tau = \tau_n \to 0$ or $\tau_n \to 1$ as the sample size n goes to infinity. Bellini et al. (2014), Mao *et al.* (2015), Bellini and Di Bernardino (2015) and Mao and Yang (2015) have already initiated and studied the connection of such extreme population expectiles with their quantile analogues when Y belongs to the domain of attraction of a Generalized Extreme Value distribution. They do not enter, however, into the crucial statistical question of how to estimate in practice these unknown tail quantities from available historical data.

In this article, we focus on high expectiles ξ_{τ_n} in the challenging maximum domain of attraction of Pareto-type distributions, where standard expectile estimates at the tails are often unstable due to data sparsity. It has been found in statistical finance and actuarial science that Pareto-type distributions describe quite well the tail structure of losses: already Embrechts *et al.* (1997, p.9) have indeed pointed out that "claims are mostly modelled by heavy-tailed distributions", and more recently Resnick (2007, p.1) has stated that "Record-breaking insurance losses, financial log-returns [...] are all examples of heavy-tailed phenomena". The rival quantile-based risk measures are investigated extensively in theoretical statistics and used widely in applied work. Notice that in applications, extreme losses correspond to tail probabilities τ_n at an extremely high level that can be even larger than $(1 - 1/n)$, see for instance Steenbergen *et al.* (2004) in the context of flood risk assessment, Embrechts and Puccetti (2007) who studied extreme operational bank losses, Cai *et al.* (2015) for an application to extreme loss returns of banks in the US market, El Methni and Stupfler (2016) who estimate several risk measures including excess-of-loss risk measures on automobile insurance data and de Valk (2016) for an application to oceanographic data. Therefore, estimating the corresponding quantile-based risk measures is a typical extreme value problem. We refer the reader to the books of Embrechts *et al.* (1997), Beirlant *et al.* (2004), and de Haan and Ferreira (2006) for a general overview of the theoretical background.

Let us point out four main conceptual results of this paper. First, we estimate the intermediate tail expectiles of order $\tau_n \to 1$ such that $n(1 - \tau_n) \to \infty$, and then extrapolate these estimates to the very extreme expectile level τ_n' which approaches one at an arbitrarily fast rate in the sense that $n(1 - \tau_n') \to c$, for some nonnegative constant c. Two such estimation methods are considered. One is indirect, based on the use of asymptotic approximations involving intermediate quantiles, and the other relies directly on least asymmetrically weighted squares (LAWS) estimation. We establish the asymptotic normality of the thus obtained estimators, which makes statistical inference for the tail expectile-based VaR feasible. Second, we wish to further contribute to the expanding literature on ES by developing a novel expectile-based variant. Taylor (2008) has already introduced an alternative expectile-based Tail Conditional Expectation (see (15) below). In contrast to his proposal, our formulation of the expectile-based ES induces a coherent risk measure. We propose two different estimators for this new coherent measure, at an extreme expectile level τ_n', and derive their full asymptotic properties. Third, we provide adapted extreme expectile-based tools for the estimation of the MES, an important factor when measuring the systemic risk of financial institutions. Denoting by X and Y, respectively, the loss of the equity return of a financial firm and that of the entire market, the MES is equal to $\mathbb{E}(X|Y > t)$, where t is a high threshold reflecting a systemic crisis, *i.e.*, a substantial market decline. For an extreme expectile $t = \xi_{\tau_n'}$ and for a wide nonparametric class of bivariate distributions of (X, Y), we construct two asymptotically normal estimators of the MES. A rival procedure by Cai *et al.* (2015) is based on extreme quantiles. Finally, we unravel the important question of how to select theoretically the extreme level τ_n' so that each expectile-based risk measure (VaR, ES, MES) at this level coincides with its quantile-based analogue at a given tail probability $\alpha_n \to 1$ as $n \to \infty$. The obtained $\tau_n' = \tau_n'(\alpha_n)$ needs itself to be estimated, which results in two final composite estimators of the risk measure. We also elucidate the asymptotic distributions of the thus built composite estimators. To our knowledge, this is the first work to actually join together the expectile perspective with the tail restrictions of extreme-value theory.

We organize this paper as follows. Section 2 discusses the basic properties of the expectile-based VaR including its connection with the standard quantile-VaR for high levels $\tau_n \to 1$. Section 3 presents the two estimation methods of intermediate and extreme expectiles. Section 4 explores the notion of expectile-based ES and discusses interesting axiomatic and asymptotic developments. Section 5 considers the problem of estimating the MES when the related variable is extreme. Section 6 addresses the important question of how to select the extreme expectile level in the three studied risk measures. The good performance of the presented procedures is shown in Section 7 and concrete applications to medical insurance data and the loss-returns of three large US investment banks are provided in Section 8.

2 Basic properties

In this paper, the generic financial position Y is a real-valued random variable, and the available data $\{Y_1, Y_2, \ldots\}$ are the negative of a series of financial returns. This implies that the right-tail of the distribution of Y corresponds to the negative of extreme losses. Following Newey and Powell (1987), the expectile ξ_τ of order $\tau \in (0, 1)$ of the random variable Y can be defined as the minimizer (1) of a piecewise-quadratic loss function or, equivalently, as

$$\xi_\tau = \mathrm{argmin}_{\theta\in\mathbb{R}} \left\{\tau\mathbb{E}\left[(Y-\theta)_+^2 - Y_+^2\right] + (1-\tau)\mathbb{E}\left[(Y-\theta)_-^2 - Y_-^2\right]\right\},$$

where $y_+ := \max(y, 0)$ and $y_- := \max(-y, 0)$. The presence of terms Y_+^2 and Y_-^2 makes indeed this problem well-defined as soon as $Y \in L^1$ [*i.e.* $\mathbb{E}|Y| < \infty$]. The first-order necessary condition for optimality related to this problem can be written in several ways, one of them being

$$\xi_\tau - \mathbb{E}(Y) = \frac{2\tau - 1}{1-\tau}\,\mathbb{E}\left[(Y - \xi_\tau)_+\right].$$

This equation has a unique solution for all $Y \in L^1$. Thenceforth expectiles of a distribution function F_Y with finite absolute first moment are well-defined, and we will assume in the sequel that $\mathbb{E}|Y| < \infty$. Expectiles summarize the distribution function in much the same way that the quantiles $q_\tau := F_Y^{-1}(\tau) = \inf\{y \in \mathbb{R} : F_Y(y) \geqslant \tau\}$ do. A justification for their use to describe distributions and their tails, as well as to quantify the "riskiness" implied by the return distribution under consideration, may be based on the following collection of elementary properties [Newey and Powell (1987), Abdous and Remillard (1995) and Bellini *et al.* (2014)]:

(i) Law invariance: a continuously differentiable distribution function is uniquely defined by its class of expectiles in the sense that the laws of two integrable random variables Y and $\widetilde{Y}$, with continuous densities, are identical if and only if $\xi_{Y,\tau} = \xi_{\tilde{Y},\tau}$ for every $\tau \in (0,1)$.

(ii) Location and scale equivariance: the τth expectile of the linear transformation $\widetilde{Y} = a + bY$, where $a, b \in \mathbb{R}$, satisfies

$$\xi_{\tilde{Y},\tau} = \begin{cases} a + b\,\xi_{Y,\tau} & \text{if} \quad b > 0 \\ a + b\,\xi_{Y,1-\tau} & \text{if} \quad b \leqslant 0\,. \end{cases}$$

(iii) Constancy: if $Y = c$ with probability 1, for some constant c (*i.e.* Y is degenerate), then $\xi_{Y,\tau} = c$ for any τ.

(iv) Strict monotonicity in τ: if $\tau_1 < \tau_2$, with $\tau_1, \tau_2 \in (0,1)$, then $\xi_{\tau_1} < \xi_{\tau_2}$. Also, the function $\tau \mapsto \xi_\tau$ maps $(0,1)$ onto its range $\{y \in \mathbb{R} : 0 < F_Y(y) < 1\}$.

(v) Preserving of stochastic order: if $Y \leqslant \tilde{Y}$ with probability 1, then $\xi_{Y,\tau} \leqslant \xi_{\tilde{Y},\tau}$ for any τ.

(vi) Subadditivity: for any variables $Y, \tilde{Y} \in L^1$, $\xi_{Y+\tilde{Y},\tau} \leqslant \xi_{Y,\tau} + \xi_{\tilde{Y},\tau}$ for all $\tau \geqslant \frac{1}{2}$. Also, $\xi_{Y+\tilde{Y},\tau} \geqslant \xi_{Y,\tau} + \xi_{\tilde{Y},\tau}$ for all $\tau \leqslant \frac{1}{2}$.

(vii) Lipschitzianity w.r.t. the Wasserstein distance: for all $Y, \tilde{Y} \in L^1$ and all $\tau \in (0,1)$, it holds that $\left|\xi_{Y,\tau} - \xi_{\tilde{Y},\tau}\right| \leqslant \tilde{\tau} \cdot d_W(Y, \tilde{Y})$, where $\tilde{\tau} = \max\left\{\frac{\tau}{1-\tau}, \frac{1-\tau}{\tau}\right\}$ and

$$d_W(Y, \tilde{Y}) = \int_{-\infty}^{\infty} |F_Y(y) - F_{\tilde{Y}}(y)|dy = \int_0^1 |F_Y^{-1}(t) - F_{\tilde{Y}}^{-1}(t)|dt.$$

(viii) Sensitivity vs resistance: expectiles are very sensitive to the magnitude of extreme observations since their gross-error-sensitivity and rejection points are infinite. Whereas they are resistant to systematic rounding and grouping since their local-shift-sensitivity is bounded.

The sign convention we have chosen for values of Y as the negative of returns implies that extreme losses correspond to levels τ close to one. Only Bellini *et al.* (2014), Mao *et al.* (2015) and Mao and Yang (2015) have described what happens for large population expectiles ξ_τ and how they are linked to extreme quantiles q_τ when F_Y is attracted to the maximum domain of Pareto-type distributions with tail index $0 < \gamma < 1$. According to Bingham *et al.* (1987), such a heavy-tailed distribution function can be expressed as

$$F_Y(y) = 1 - \ell(y) \cdot y^{-1/\gamma} \tag{3}$$

where $\ell(\cdot)$ is a slowly-varying function at infinity, *i.e*, $\ell(\lambda y)/\ell(y) \to 1$ as $y \to \infty$ for all $\lambda > 0$. The tail index γ tunes the tail heaviness of the distribution function F_Y. Note also that the first moment of F_Y does not exist when $\gamma > 1$. For most applicational purposes in risk management, it has been found in previous studies that assumption (3) describes sufficiently well the tail structure of actuarial and financial data: in addition to the monographs of Embrechts *et al.* (1997) and Resnick (2007), see for instance Chavez-Demoulin *et al.* (2015) and the references therein. See also Alm (2015) for a recent study in the context of the Swedish insurance market.

Writing $\overline{F}_Y := 1 - F_Y$, Bellini *et al.* (2014) have shown in the case $\gamma < 1$ that

$$\frac{\overline{F}_Y(\xi_\tau)}{\overline{F}_Y(q_\tau)} \sim \gamma^{-1} - 1 \quad \text{as} \quad \tau \to 1, \tag{4}$$

or equivalently $\frac{\overline{F}_Y(\xi_\tau)}{1-\tau} \sim \gamma^{-1} - 1$ as $\tau \to 1$. It follows that extreme expectiles ξ_τ are larger than extreme quantiles q_τ (*i.e.* $\xi_\tau > q_\tau$) when $\gamma > \frac{1}{2}$, whereas $\xi_\tau < q_\tau$ when $\gamma < \frac{1}{2}$, for all large τ. The connection (4) between high expectiles and quantiles can actually be refined appreciably by considering a strengthened yet classical version of condition (3). Assume that the tail quantile function U of Y, namely the left-continuous inverse of $1/\overline{F}_Y$, defined by

$$\forall t > 1, \; U(t) = \inf\left\{ y \in \mathbb{R} \,\middle|\, \frac{1}{\overline{F}_Y(y)} \geqslant t \right\},$$

is such that there exist $\gamma > 0$, $\rho \leqslant 0$, and a function $A(\cdot)$ converging to 0 at infinity and having constant sign such that

$\boldsymbol{\mathcal{C}_2(\gamma, \rho, A)}$ For all $x > 0$,

$$\lim_{t\to\infty} \frac{1}{A(t)} \left[\frac{U(tx)}{U(t)} - x^\gamma \right] = x^\gamma \frac{x^\rho - 1}{\rho}.$$

Here and in what follows, $(x^\rho - 1)/\rho$ is to be understood as $\log x$ when $\rho = 0$. The interpretation of this extremal value condition can be found in Beirlant *et al.* (2004) and de Haan and Ferreira (2006) along with abundant examples of commonly used continuous distributions satisfying $\mathcal{C}_2(\gamma, \rho, A)$: for instance, the (Generalized) Pareto, Burr, Fréchet, Student, Fisher and Inverse-Gamma distributions all satisfy this condition, and more generally so does any distribution whose distribution function F satisfies

$$1 - F(x) = x^{-1/\gamma}\left(a + bx^{\rho/\gamma} + \mathrm{o}(x^{\rho/\gamma})\right) \quad \text{as} \quad x \to \infty,$$

where a and b are positive constants and $\rho < 0$. This contains in particular the Hall-Weiss class of models (see Hua and Joe, 2011).

Proposition 1. *Assume that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds, with $0 < \gamma < 1$. Then*

$$\begin{aligned} \frac{\overline{F}_Y(\xi_\tau)}{1-\tau} &= (\gamma^{-1} - 1)(1 + \varepsilon(\tau)) \\ \textit{with } \varepsilon(\tau) &= -\frac{(\gamma^{-1}-1)^\gamma \mathbb{E}(Y)}{q_\tau}(1 + \mathrm{o}(1)) - \frac{(\gamma^{-1}-1)^{-\rho}}{\gamma(1-\rho-\gamma)} A((1-\tau)^{-1})(1 + \mathrm{o}(1)) \ \textit{as } \tau \uparrow 1. \end{aligned}$$

Even more strongly, one can establish the precise bias term in the asymptotic approximation of (ξ_τ / q_τ) itself.

Corollary 1. *Assume that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds, with $0 < \gamma < 1$. If F_Y is strictly increasing, then*

$$\begin{aligned} \frac{\xi_\tau}{q_\tau} &= (\gamma^{-1} - 1)^{-\gamma}(1 + r(\tau)) \\ \textit{with } r(\tau) &= \frac{\gamma(\gamma^{-1}-1)^\gamma \mathbb{E}(Y)}{q_\tau}(1 + \mathrm{o}(1)) \\ &\quad + \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\rho-\gamma} + \frac{(\gamma^{-1}-1)^{-\rho} - 1}{\rho} + \mathrm{o}(1)\right) A((1-\tau)^{-1}) \ \textit{as } \tau \uparrow 1. \end{aligned}$$

Other refinements under similar second order regular variation conditions can also be found in Mao *et al.* (2015) and Mao and Yang (2015). An extension to a subset of the challenging Gumbel domain of attraction is also derived in Proposition 2.4 in Bellini and Di Bernardino (2015). In practice, the tail quantities ξ_τ, q_τ and γ are unknown and only a sample of random copies $(Y_1, \ldots, Y_n)$ of Y is typically available. While extreme-value estimates of high quantiles and of the tail index γ are used widely in applied work and investigated extensively in theoretical statistics, the problem of estimating ξ_τ, when $\tau = \tau_n \to 1$ at an arbitrary rate as $n \to \infty$, has not been addressed yet. Direct expectile estimates at the tails are incapable of extrapolating outside the data and are often unstable due to data sparseness. This motivated us to construct estimators of large expectiles ξ_{τ_n} and derive their limit distributions when they are located in the range of the data or near and even beyond the sample maximum. We shall assume the extended regular variation condition $\mathcal{C}_2(\gamma, \rho, A)$ to obtain some convergence results.

3 Estimation of the expectile-based VaR

Our main objective in this section is to estimate ξ_{τ_n} for high levels τ_n that may approach one at any rate, covering both scenarios of intermediate expectiles with $n(1-\tau_n) \to \infty$ and extreme expectiles with $n(1-\tau_n) \to c$, for some nonnegative constant c. We assume that the available data consists of an n-tuple $(Y_1, \ldots, Y_n)$ of independent copies of Y, and denote by $Y_{1,n} \leqslant \cdots \leqslant Y_{n,n}$ their ascending order statistics.

3.1 Intermediate expectile estimation

Here, we first use an indirect estimation method based on intermediate quantiles, and then discuss a direct asymmetric least squares estimation method.

3.1.1 Estimation based on intermediate quantiles

The rationale for this first method relies on the regular variation property (3) and on the asymptotic equivalence (4). Given that $\overline{F}_Y$ is regularly varying at infinity with index $-1/\gamma$ [*i.e.* it satisfies, for any $x > 0$, the property $\overline{F}_Y(tx)/\overline{F}_Y(t) \to x^{-1/\gamma}$ as $t \to \infty$], it follows that U is regularly varying as well with index γ, see *e.g.* Proposition B.1.9.9 in de Haan and Ferreira (2006). Hence, (4) entails that

$$\frac{\xi_\tau}{q_\tau} \sim (\gamma^{-1} - 1)^{-\gamma} \quad \text{as} \quad \tau \uparrow 1. \tag{5}$$

This result is also an immediate consequence of Corollary 1 above and can be found in Proposition 2.3 of Bellini and Di Bernardino (2015) as well. Therefore, for a suitable estimator $\widehat{\gamma}$ of γ, we may suggest estimating the intermediate expectile ξ_{τ_n} by

$$\widehat{\xi}_{\tau_n} := (\widehat{\gamma}^{-1} - 1)^{-\widehat{\gamma}}\, \widehat{q}_{\tau_n}, \quad \text{where} \quad \widehat{q}_{\tau_n} := Y_{n-\lfloor n(1-\tau_n)\rfloor,n}$$

and $\lfloor \cdot \rfloor$ stands for the floor function. This estimator parallels the intermediate quantile-VaR $\widehat{q}_{\tau_n}$ and crucially hinges on the estimated tail index $\widehat{\gamma}$. Accordingly, it is more extreme than $\widehat{q}_{\tau_n}$ when $\widehat{\gamma} > \frac{1}{2}$, but less extreme when $\widehat{\gamma} < \frac{1}{2}$. A simple and widely used estimator of γ is given by the popular Hill estimator (Hill, 1975):

$$\widehat{\gamma}_H = \frac{1}{k} \sum_{i=1}^{k} \log \frac{Y_{n-i+1,n}}{Y_{n-k,n}}, \tag{6}$$

where $k = k(n)$ is an intermediate sequence in the sense that $k(n) \to \infty$ such that $k(n)/n \to 0$ as $n \to \infty$. The monographs of Beirlant *et al.* (2004) and de Haan and Ferreira (2006) give a nice summary of the properties of $\widehat{\gamma}_H$ and review other efficient estimation methods with an extensive bibliography.

Next, we formulate conditions that lead to asymptotic normality for $\widehat{\xi}_{\tau_n}$.

Theorem 1. *Assume that F_Y is strictly increasing, that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds with $0 < \gamma < 1$, that $\tau_n \uparrow 1$ and $n(1-\tau_n) \to \infty$. Assume further that*

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma} - \gamma, \frac{\widehat{q}_{\tau_n}}{q_{\tau_n}} - 1\right) \xrightarrow{d} (\Gamma, \Theta). \tag{7}$$

If $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$, then

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \xrightarrow{d} m(\gamma)\Gamma + \Theta - \lambda$$

with $m(\gamma) := (1-\gamma)^{-1} - \log(\gamma^{-1} - 1)$ and

$$\lambda := \gamma(\gamma^{-1} - 1)^{\gamma}\mathbb{E}(Y)\lambda_1 + \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\rho-\gamma} + \frac{(\gamma^{-1}-1)^{-\rho} - 1}{\rho}\right)\lambda_2.$$

When using the Hill estimator (6) of γ with $k = \lceil n(1-\tau_n) \rceil$, sufficient regularity conditions for (7) to hold can be found in Theorems 2.4.1 and 3.2.5 in de Haan and Ferreira (2006, p.50 and p.74). Under these conditions, the limit distribution Γ is then Gaussian with mean $\lambda_2/(1-\rho)$ and variance γ^2, while Θ is the centered Gaussian distribution with variance γ^2. Lemma 3.2.3 in de Haan and Ferreira (2006, p.71) shows that both Gaussian limiting distributions are independent. As an immediate consequence we get the following for $\widehat{\gamma} = \widehat{\gamma}_H$.

Corollary 2. *Assume that F_Y is strictly increasing, that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds with $0 < \gamma < 1$, that $\tau_n \uparrow 1$ and $n(1-\tau_n) \to \infty$. If $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$, then*

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \xrightarrow{d} \mathcal{N}\left(\frac{m(\gamma)}{1-\rho}\lambda_2 - \lambda, v(\gamma)\right),$$

with $m(\gamma)$ and λ as in Theorem 1, and

$$v(\gamma) = \gamma^2\left[1 + \left(\frac{1}{1-\gamma} - \log\left(\frac{1}{\gamma} - 1\right)\right)^2\right].$$

Yet, a drawback to the resulting estimator $\widehat{\xi}_{\tau_n}$ lies in its heavy dependency on the estimated quantile $\widehat{q}_{\tau_n}$ and tail index $\widehat{\gamma}$ in the sense that $\widehat{\xi}_{\tau_n}$ may inherit the vexing defects of both $\widehat{q}_{\tau_n}$ and $\widehat{\gamma}$. Note also that $\widehat{\xi}_{\tau_n}$ is asymptotically biased, which is not the case for $\widehat{q}_{\tau_n}$; it should be pointed out though that one may design a bias-reduced version of the estimator $\widehat{\xi}_{\tau_n}$. Indeed, the bias components λ_1 and λ_2 appearing in Theorem 1 can be estimated, respectively, by using $\widehat{\lambda}_1 = \sqrt{n(1-\tau_n)}\widehat{q}_{\tau_n}^{-1}$ and by applying the methodology of Caeiro *et al.* (2005) in conjunction with the Hall-Welsh class of models to get an estimator $\widehat{\lambda}_2$ of

λ_2. Along with the empirical mean $\overline{Y}$, the estimator $\widehat{\gamma}$, and a consistent estimator $\widehat{\rho}$ of the second-order parameter ρ (a review of possible estimators $\widehat{\rho}$ is given in Gomes and Guillou, 2015), it is possible to come up with a consistent estimator

$$\widehat{\lambda} := \widehat{\gamma}(\widehat{\gamma}^{-1}-1)^{\widehat{\gamma}}\overline{Y}\widehat{\lambda}_1 + \left(\frac{(\widehat{\gamma}^{-1}-1)^{-\widehat{\rho}}}{1-\widehat{\rho}-\widehat{\gamma}} + \frac{(\widehat{\gamma}^{-1}-1)^{-\widehat{\rho}}-1}{\widehat{\rho}}\right)\widehat{\lambda}_2$$

of the bias component λ. This in turn enables one to define a bias-reduced version of $\widehat{\xi}_{\tau_n}$, for instance, as

$$\widehat{\xi}_{\tau_n}^{RB} := \widehat{\xi}_{\tau_n}\left(1 - \left[\frac{m(\widehat{\gamma})}{1-\widehat{\rho}}\widehat{\lambda}_2 - \widehat{\lambda}\right]\frac{1}{\sqrt{n(1-\tau_n)}}\right).$$

Of course, one should expect the value of the asymptotic variance of this estimator to be even higher than that of $\widehat{\xi}_{\tau_n}$, as when bias reduction techniques are applied to the Hill estimator (see *e.g.* Theorem 3.2 in Caeiro *et al.*, 2005).

Another efficient way of estimating ξ_{τ_n}, which we develop in the next section, is by joining together the least asymmetrically weighted squares (LAWS) estimation with the tail restrictions of modern extreme-value theory.

3.1.2 Asymmetric least squares estimation

Here, we consider estimating the expectile ξ_{τ_n} by its empirical counterpart defined through

$$\widetilde{\xi}_{\tau_n} = \arg\min_{u\in\mathbb{R}} \frac{1}{n}\sum_{i=1}^{n} \eta_{\tau_n}(Y_i - u),$$

where $\eta_\tau(y) = |\tau - 1\!\!1\{y \leqslant 0\}|y^2$ is the expectile check function. This LAWS minimizer can easily be calculated by applying the function *"expectile"* implemented in the R package *'expectreg'*. It is not hard to verify that

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) = \arg\min_{u\in\mathbb{R}} \psi_n(u) \tag{8}$$

$$\text{with} \qquad \psi_n(u) := \frac{1}{2\xi_{\tau_n}^2}\sum_{i=1}^{n}\left[\eta_{\tau_n}(Y_i - \xi_{\tau_n} - u\xi_{\tau_n}/\sqrt{n(1-\tau_n)}) - \eta_{\tau_n}(Y_i - \xi_{\tau_n})\right].$$

It follows from the continuity and the convexity of η_τ that (ψ_n) is a sequence of almost surely continuous and convex random functions. A result of Geyer (1996) [see also Theorem 5 in Knight (1999)] then states that to examine the convergence of the left-hand side term of (8), it is enough to investigate the asymptotic properties of the sequence (ψ_n). Built on this idea, we get the asymptotic normality of the LAWS estimator $\widetilde{\xi}_{\tau_n}$ by applying standard techniques involving sums of independent and identically distributed random variables. We will denote in the sequel by Y_- the negative part of Y, *i.e.*, $Y_- = \min(Y, 0)$.

Theorem 2. *Assume that there is $\delta > 0$ such that $\mathbb{E}|Y_-|^{2+\delta} < \infty$, that $0 < \gamma < 1/2$ and $\tau_n \uparrow 1$ is such that $n(1-\tau_n) \to \infty$. Then*

$$\sqrt{n(1-\tau_n)}\left(\frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, V(\gamma)\right) \quad \text{with} \quad V(\gamma) = \frac{2\gamma^3}{1-2\gamma}.$$

Interestingly, in contrast to Theorem 1 and Corollary 2, the limit distribution in Theorem 2 is derived without recourse to either the extended regular variation condition $\mathcal{C}_2(\gamma, \rho, A)$ or any bias condition. A mild moment assumption and the condition $0 < \gamma < 1/2$ suffice. It has been found in previous studies by many authors (*e.g.* recently by Chavez-Demoulin *et al.* (2014), Alm (2015), Cai *et al.* (2015) and El Methni and Stupfler (2016)) that the model assumption of Pareto-type tails along with these conditions (required for losses to have at least a finite variance) deliver competitive results for most applicational purposes in risk management. In these studies the realized values of γ were found below 1/2, being in line with the recent findings of Cai *et al.* (2015) and our findings in Section 8. We also refer the reader to the R package 'CASdatasets' which contains a large variety of dataset examples where realized values of γ often vary between 1/4 and 1/2. Most importantly, unlike the indirect expectile estimator $\widehat{\xi}_{\tau_n}$, the new estimator $\widetilde{\xi}_{\tau_n}$ does not hinge by construction on any particular type of quantile or tail index estimators. A comparison of the asymptotic variance $V(\gamma)$ of $\widetilde{\xi}_{\tau_n}$ with the asymptotic variance $v(\gamma)$ of $\widehat{\xi}_{\tau_n}$ is provided in Figure 1. It can be seen that both asymptotic variances are extremely stable and close for values of $\gamma < 0.3$, with a slight advantage for $V(\gamma)$ in dashed line. Then $V(\gamma)$ becomes appreciably larger than $v(\gamma)$ for $\gamma > 0.3$ and explodes in the neighborhood of 1/2, while $v(\gamma)$ in solid line remains lower than the level 1.25.

3.2 Extreme expectile estimation

We now discuss the important issue of estimating extreme tail expectiles $\xi_{\tau_n'}$, where $\tau_n' \uparrow 1$ with $n(1-\tau_n') \to c < \infty$ as $n \to \infty$. The basic idea is to extrapolate intermediate expectile estimates of order $\tau_n \to 1$, such that $n(1-\tau_n) \to \infty$, to the very extreme level τ_n'. This is achieved by transferring the elegant device of Weissman (1978) for estimating an extreme quantile to our expectile setup. Note that, in standard extreme-value theory and related fields of application, the levels τ_n' and τ_n are typically set to be $\tau_n' = 1 - p_n$ for a p_n much smaller than $\frac{1}{n}$, and $\tau_n = 1 - \frac{k(n)}{n}$ for an intermediate sequence of integers $k(n)$.

The model assumption of Pareto-type tails (3) means that $U(tx)/U(t) \to x^\gamma$ as $t \to \infty$, which in turn suggests that

$$\frac{q_{\tau_n'}}{q_{\tau_n}} = \frac{U((1-\tau_n')^{-1})}{U((1-\tau_n)^{-1})} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma} \quad \text{and thus} \quad \frac{\xi_{\tau_n'}}{\xi_{\tau_n}} \approx \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\gamma}$$

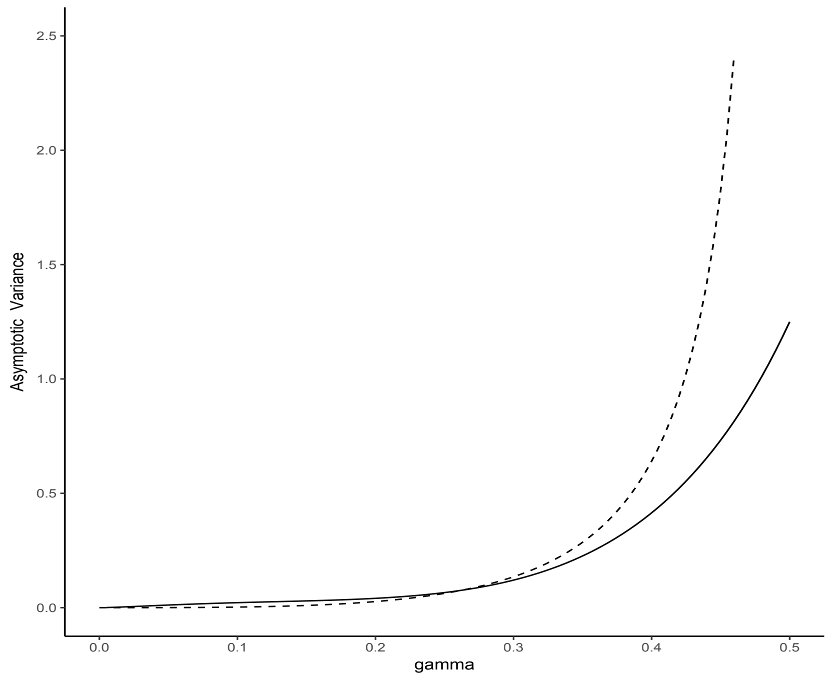

Figure 1: *Asymptotic variances $V(\gamma)$ of the LAWS estimator $\widetilde{\xi}_{\tau_n}$ in dashed line and $v(\gamma)$ of the indirect estimator $\widehat{\xi}_{\tau_n}$ in solid line, with $\gamma \in (0, 1/2)$.*

by (5), for τ_n, τ_n' satisfying suitable conditions. This approximation motivates the following class of plug-in estimators of $\xi_{\tau_n'}$:

$$\overline{\xi}^{\star}_{\tau_n'} \equiv \overline{\xi}^{\star}_{\tau_n'}(\tau_n) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}} \overline{\xi}_{\tau_n} \tag{9}$$

where $\widehat{\gamma}$ is an estimator of γ, and $\overline{\xi}_{\tau_n}$ stands for either the estimator $\widehat{\xi}_{\tau_n}$ or $\widetilde{\xi}_{\tau_n}$ of the intermediate expectile ξ_{τ_n}. As a matter of fact, we have $\overline{\xi}^{\star}_{\tau_n'}/\overline{\xi}_{\tau_n} = \widehat{q}^{\star}_{\tau_n'}/\widehat{q}_{\tau_n}$ where $\widehat{q}_{\tau_n} = Y_{n-\lfloor n(1-\tau_n)\rfloor, n}$ is the intermediate quantile estimator introduced above, and $\widehat{q}^{\star}_{\tau_n'}$ is the extreme Weissman quantile estimator defined as

$$\widehat{q}^{\star}_{\tau_n'} \equiv \widehat{q}^{\star}_{\tau_n'}(\tau_n) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}} \widehat{q}_{\tau_n}. \tag{10}$$

We then show that $(\frac{\overline{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}} - 1)$ has the same limit distribution as $(\widehat{\gamma} - \gamma)$ with a different scaling.

Theorem 3. *Assume that F_Y is strictly increasing, that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds with $\rho < 0$, that τ_n, $\tau_n' \uparrow 1$, with $n(1-\tau_n) \to \infty$ and $n(1-\tau_n') \to c < \infty$. If moreover*

$$\sqrt{n(1-\tau_n)}\left(\frac{\overline{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) \xrightarrow{d} \Delta \quad \text{and} \quad \sqrt{n(1-\tau_n)}(\widehat{\gamma} - \gamma) \xrightarrow{d} \Gamma,$$

with $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/\log[(1-$

$\tau_n)/(1-\tau_n')] \to \infty$, *then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\overline{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}-1\right) \xrightarrow{d} \Gamma.$$

More specifically, we can choose $\overline{\xi}_{\tau_n}$ in (9) to be either the indirect intermediate expectile estimator $\widehat{\xi}_{\tau_n}$, the resulting extreme expectile estimator $\widehat{\xi}^{\star}_{\tau_n'} := \overline{\xi}^{\star}_{\tau_n'}$ being

$$\widehat{\xi}^{\star}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}} \widehat{\xi}_{\tau_n} = \left(\widehat{\gamma}^{-1}-1\right)^{-\widehat{\gamma}} \widehat{q}^{\star}_{\tau_n'}, \tag{11}$$

or we may choose $\overline{\xi}_{\tau_n}$ to be the LAWS estimator $\widetilde{\xi}_{\tau_n}$, yielding the extreme expectile estimator

$$\widetilde{\xi}^{\star}_{\tau_n'} = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}} \widetilde{\xi}_{\tau_n}, \tag{12}$$

Their respective asymptotic properties are given in the next two corollaries of Theorem 3.

Corollary 3. *Assume that F_Y is strictly increasing, that condition $\mathcal{C}_2(\gamma,\rho,A)$ holds with $0<\gamma<1$ and $\rho<0$, and that τ_n, $\tau_n' \uparrow 1$ with $n(1-\tau_n)\to\infty$ and $n(1-\tau_n')\to c<\infty$. Assume further that*

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}-\gamma, \frac{\widehat{q}_{\tau_n}}{q_{\tau_n}}-1\right) \xrightarrow{d} (\Gamma,\Theta).$$

If $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau_n')] \to \infty$, then

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widehat{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}-1\right) \xrightarrow{d} \Gamma.$$

Corollary 4. *Assume that F_Y is strictly increasing, there is $\delta>0$ such that $\mathbb{E}|Y_-|^{2+\delta}<\infty$, condition $\mathcal{C}_2(\gamma,\rho,A)$ holds with $0<\gamma<1/2$ and $\rho<0$, and that τ_n, $\tau_n'\uparrow 1$ with $n(1-\tau_n)\to\infty$ and $n(1-\tau_n')\to c<\infty$. If in addition*

$$\sqrt{n(1-\tau_n)}(\widehat{\gamma}-\gamma) \xrightarrow{d} \Gamma$$

and $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau_n')] \to \infty$, then

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widetilde{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}-1\right) \xrightarrow{d} \Gamma.$$

4 Expectile-based expected shortfall

The conventional quantile-based VaR was often criticized for being insensitive to the magnitude of extreme losses since it only depends on the frequency of tail losses and not on their values. Acerbi (2002), Rockafellar and Uryasev (2002) proposed to change the measurement method for calculating losses from the usual quantile-VaR to an alternative coherent method known as Expected Shortfall (ES). This proposal was criticized though for its dependency only on the tail event. The formulation of the ES remains still intrinsically linked to quantiles as can be seen from (13) and (14) below. This motivated Kuan *et al.* (2009) to introduce the expectile-based VaR which depends on both the tail realizations of the loss variable and their probability. Yet, with the recent crisis in the financial industry, the vast majority of market participants (investors, risk managers, clearing houses), academics and regulators are more concerned with the risk exposure to a catastrophic event that might wipe out an investment in terms of the size of potential losses. Already in October 2013, The Basel Committee on Banking Supervision proposed to change the traditional VaR with the ES for calculating losses. This motivates us to introduce a new variant of ES which is purely built on expectiles, inherits their coherency, but is more alert to extreme risks.

4.1 Basic properties

The standard ES, also known under the names Conditional Value at Risk or Average Value at Risk, is defined as the average of the quantile function above a given confidence level τ. It is traditionally expressed at the $100(1-\tau)\%$ security level as the Quantile-based Expected Shortfall (QES):

$$\mathrm{QES}(\tau) := \frac{1}{1-\tau}\int_\tau^1 q_\alpha d\alpha. \tag{13}$$

When the financial position Y is continuous, $\mathrm{QES}(\tau)$ is just the conditional expectation

$$\mathrm{QTCE}(\tau) := \mathbb{E}[Y|Y > q_\tau] \tag{14}$$

of Y given that it exceeds the VaR q_τ. In this sense, it is referred to as Tail Conditional Expectation (TCE), with $-\mathrm{QES}(\tau)$ being interpreted as the expected return on the portfolio in the worst $100(1-\tau)\%$ of cases. Similarly, Taylor (2008) has suggested to use an expectile-based TCE defined as

$$\begin{aligned} \mathrm{XTCE}(\tau) &:= \mathbb{E}[Y|Y > \xi_\tau] \qquad (15)\\ &= \left(1+\frac{1-\tau}{(2\tau-1)\overline{F}_Y(\xi_\tau)}\right)\xi_\tau - \frac{1-\tau}{(2\tau-1)\overline{F}_Y(\xi_\tau)}\mathbb{E}(Y). \end{aligned}$$

We shall discuss below that this proposal does not seem, however, to be a coherent risk measure in general. Instead, we introduce the alternative expectile-based ES

$$\mathrm{XES}(\tau) := \frac{1}{1-\tau}\int_{\tau}^{1} \xi_{\alpha} d\alpha, \tag{16}$$

which defines a new coherent risk measure as established in Proposition 4 below. We also show in Proposition 2, under the model assumption of Pareto-type distributions $F_Y(\cdot)$ with tail index $\gamma < 1$, that $\mathrm{XES}(\tau)$ is asymptotically equivalent to $\mathrm{XTCE}(\tau)$ as $\tau \to 1$, and hence inherits its direct meaning as a conditional expectation for all τ large enough. On the other hand, the choice between the expectile-based ES/TCE and their quantile-based versions will depend on the value at hand of $\gamma \lesseqgtr \frac{1}{2}$ as is the case in the duality between the expectile-based VaR and quantile-VaR. More precisely, the $\mathrm{XES}(\tau)$ in (16) and $\mathrm{XTCE}(\tau)$ in (15) are more extreme (respectively, less extreme) than their quantile-based analogues $\mathrm{QES}(\tau)$ in (13) and $\mathrm{QTCE}(\tau)$ in (14), for all τ large enough, when $\gamma > \frac{1}{2}$ (respectively, $\gamma < \frac{1}{2}$).

Proposition 2. *Assume that the distribution of Y belongs to the Fréchet maximum domain of attraction with tail index $\gamma < 1$, or equivalently, that condition (3) holds. Then*

$$\frac{XES(\tau)}{QES(\tau)} \sim \frac{\xi_\tau}{q_\tau} \sim \frac{XTCE(\tau)}{QTCE(\tau)} \quad \text{and} \quad \frac{XES(\tau)}{\xi_\tau} \sim \frac{1}{1-\gamma} \sim \frac{XTCE(\tau)}{\xi_\tau} \quad \text{as} \quad \tau \to 1.$$

These connections are very useful when it comes to proposing estimators for $\mathrm{XES}(\tau)$ and $\mathrm{XTCE}(\tau)$. One may also establish, in the spirit of Proposition 1, a precise control of the remainder term which arises when using Proposition 2. This will prove to be quite useful when examining the asymptotic properties of the extreme expectile-based ES estimators.

Proposition 3. *Assume that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds, with $0 < \gamma < 1$. Then, as $\tau \to 1$,*

$$\begin{aligned}
\frac{XES(\tau)}{\xi_\tau} &= \frac{1}{1-\gamma}\Bigg(1 - \frac{\gamma^2(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y)}{q_\tau}(1+\mathrm{o}(1)) \\
&\quad + \frac{1-\gamma}{(1-\rho-\gamma)^2}(\gamma^{-1}-1)^{-\rho}A((1-\tau)^{-1})(1+\mathrm{o}(1))\Bigg), \\
\frac{XTCE(\tau)}{\xi_\tau} &= \frac{1}{1-\gamma}\left[1 + \frac{(\gamma^{-1}-1)^{-\rho}}{1-\rho-\gamma}A((1-\tau)^{-1})(1+\mathrm{o}(1)) + \mathrm{o}(q_\tau^{-1})\right].
\end{aligned}$$

From the point of view of the axiomatic theory, an influential paper in the literature by Artzner *et al.* (1999) provides an axiomatic foundation for coherent risk measures. Like the quantile-based expected shortfall $\mathrm{QES}(\tau)$ and the expectile-based VaR, we show that the expectile-based expected shortfall $\mathrm{XES}(\tau)$ satisfies all of their requirements, namely Translation invariance, Positive homogeneity, Monotonicity and Subadditivity. However, while the Tail Conditional Expectation variant $\mathrm{XTCE}(\tau)$ can easily be shown to be translation invariant and positive homogeneous, its monotonicity and subadditivity remain unclear. It should also be noted that, in contrast to QES(τ), the coherence of XES(τ) is actually a straightforward consequence of the coherence of the expectile-based VaR above the median in conjunction with the fact that the expectile-based ES is an increasing linear functional of the expectile-based VaR above some high level, in the sense that

$$\xi_{\alpha}^{(1)} \leqslant \xi_{\alpha}^{(2)} \ \forall \alpha \in (\tau, 1) \ \Rightarrow \ \text{XES}^{(1)}(\tau) = \frac{1}{1-\tau}\int_{\tau}^{1} \xi_{\alpha}^{(1)} d\alpha \leqslant \frac{1}{1-\tau}\int_{\tau}^{1} \xi_{\alpha}^{(2)} d\alpha = \text{XES}^{(2)}(\tau).$$

Proposition 4. *For all $\tau \geqslant 1/2$, the expectile-based expected shortfall XES(τ) induces a coherent risk measure, and the tail conditional expectation XTCE(τ) is translation invariant and positive homogeneous.*

This result does not seem to have been appreciated in the literature before. It affords an additional convincing reason that the use of both expectile-based VaR and ES may be preferred over the classical quantile-based versions.

4.2 Estimation and asymptotics

When analyzing the very far tails of the involved distribution, as required in modern regulatory frameworks (such as the European Union Solvency II directive), financial institutions and insurance companies are typically interested in the region $\tau = \tau_n' \uparrow 1$, as the sample size $n \to \infty$. This is particularly required to manage extreme events. For example, Acharya *et al.* (2012) handle once-in-a-decade events with one year of data. Gilli and Këllezi (2006) apply extreme value theory to estimate risk measures for several stock market indices. Hartmann *et al.* (2004, 2005, 2010) employ extreme value techniques to evaluate systemic risk in the European and American banking systems. Also, Ibragimov *et al.* (2009) develop models for catastrophe insurance markets based on real data. The asymptotic equivalence XES$(\tau_n') \sim (1-\gamma)^{-1}\xi_{\tau_n'}$, established in Proposition 2, suggests the following estimators of the expectile-based ES:

$$\widehat{\text{XES}}^{\star}(\tau_n') = (1-\widehat{\gamma})^{-1} \cdot \widehat{\xi}_{\tau_n'}^{\star} \quad \text{and} \quad \widetilde{\text{XES}}^{\star}(\tau_n') = (1-\widehat{\gamma})^{-1} \cdot \widetilde{\xi}_{\tau_n'}^{\star} \tag{17}$$

where $\widehat{\xi}_{\tau_n'}^{\star}$ and $\widetilde{\xi}_{\tau_n'}^{\star}$ are the extreme expectile estimators defined above in (11)-(12), and $\widehat{\gamma}$ is an estimator of γ. Another option motivated by the second asymptotic equivalence XES$(\tau_n') \sim \frac{\xi_{\tau_n'}}{q_{\tau_n'}} \cdot$ QES(τ_n') would be to estimate XES(τ_n') by

$$\widehat{\text{XES}}^{\dagger}(\tau_n') = \widehat{\xi}_{\tau_n'}^{\star} \cdot \frac{\widehat{\text{QES}}^{\star}(\tau_n')}{\widehat{q}_{\tau_n'}^{\star}} \quad \text{or} \quad \widetilde{\text{XES}}^{\dagger}(\tau_n') = \widetilde{\xi}_{\tau_n'}^{\star} \cdot \frac{\widehat{\text{QES}}^{\star}(\tau_n')}{\widehat{q}_{\tau_n'}^{\star}} \tag{18}$$

for a suitable estimator $\widehat{\text{QES}}^{\star}(\tau_n')$ of QES(τ_n') (see, *e.g.*, El Methni *et al.* (2014)), with $\widehat{q}_{\tau_n'}^{\star}$ being the extreme Weissman quantile estimator defined in (10). Our experience with real and simulated data indicates, however, that the estimates $\widehat{\mathrm{XES}}^{\star}(\tau_n')$ and $\widehat{\mathrm{XES}}^{\dagger}(\tau_n')$ [respectively, $\widetilde{\mathrm{XES}}^{\star}(\tau_n')$ and $\widetilde{\mathrm{XES}}^{\dagger}(\tau_n')$] point toward very similar results. We therefore restrict our theoretical treatment to the initial versions given in (17). Our first asymptotic result is for the extreme XES estimator $\widehat{\mathrm{XES}}^{\star}(\tau_n')$:

Corollary 5. *Under the conditions of Corollary 3,*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widehat{XES}^{\star}(\tau_n')}{XES(\tau_n')}-1\right)\xrightarrow{d}\Gamma.$$

In what concerns the asymmetric least squares-type estimator $\widetilde{\mathrm{XES}}^{\star}(\tau_n')$, we have the following result.

Corollary 6. *Under the conditions of Corollary 4,*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widetilde{XES}^{\star}(\tau_n')}{XES(\tau_n')}-1\right)\xrightarrow{d}\Gamma.$$

Both results are derived by noticing that, on the one hand, the extreme expectile estimators $\widehat{\xi}^{\star}_{\tau_n'}$ and $\widetilde{\xi}^{\star}_{\tau_n'}$ converge to the same distribution as the estimator $\widehat{\gamma}$ but with a slower rate in view of Corollaries 3 and 4. On the other hand, the nonrandom remainder term coming from the use of Proposition 2 can be controlled by applying Proposition 3, so detailed proofs are omitted.

When it comes to estimate the Tail Conditional Expectation $\mathrm{XTCE}(\tau_n')$, the asymptotic equivalence $\mathrm{XTCE}(\tau_n') \sim \mathrm{XES}(\tau_n')$ as $n \to \infty$, obtained in Proposition 2, suggests to use the same estimators of $\mathrm{XES}(\tau_n')$ in (17) and (18) for $\mathrm{XTCE}(\tau_n')$ itself. To derive the limit distributions here, the basic arguments go as above.

5 Marginal expected shortfall

5.1 Setting and objective

With the recent financial crisis and the rising interconnection between financial institutions, interest in the concept of systemic risk has grown. Acharya *et al.* (2012), Brownlees and Engle (2016) and Engle *et al.* (2015) define systemic risk as the propensity of a financial institution to be undercapitalized when the financial system as a whole is undercapitalized. They have proposed econometric and statistical approaches to measure the systemic risk of financial institutions. An important step in constructing a systemic risk measure for a financial firm is to measure the contribution of the firm to a systemic crisis. A systemic event or crisis is specified as a major stock market decline that happens once or twice a decade. The total risk measured by the expected capital shortfall in the financial system during a systemic crisis is typically decomposed into firm level contributions. Each financial firm's contribution to systemic risk can then be measured as its marginal expected shortfall (MES), *i.e.*, the expected loss on its equity return conditional on the occurrence of an extreme loss in the aggregated return of the financial market. More specifically, denote the loss return on the equity of a financial firm as X and that of the entire market as Y. Then the MES at probability level $(1-\tau)$ is defined as

$$\text{QMES}(\tau) = \mathbb{E}\{X|Y > q_{Y,\tau}\}, \quad \tau \in (0,1),$$

where $q_{Y,\tau}$ is the τth quantile of the distribution of Y. Typically, a systemic crisis defined as an extreme tail event corresponds to a probability τ at an extremely high level that can be even larger than $(1-1/n)$, where n is the sample size of historical data that are used for estimating QMES(τ). The estimation procedure in Acharya *et al.* (2012) relies on daily data from only 1 year and assumes a specific linear relationship between X and Y. A nonparametric kernel estimation method has been performed in Brownlees and Engle (2016) and Engle *et al.* (2015), but cannot handle extreme events required for systemic risk measures (*i.e.* $1-\tau = \text{O}(1/n)$). Very recently, Cai *et al.* (2015) have proposed adapted extreme-value tools for the estimation of QMES(τ) without recourse to any parametric structure on (X,Y). Here, instead of the extreme τth quantile $q_{Y,\tau}$, we will explore the use of the τth expectile analogue $\xi_{Y,\tau}$ in the marginal expected shortfall

$$\text{XMES}(\tau) = \mathbb{E}\{X|Y > \xi_{Y,\tau}\}$$

at least for the following reason: as claimed by Newey and Powell (1987), Kuan *et al.* (2009) and Sobotka and Kneib (2012) among others, expectiles make a more efficient use of the available data since they rely on the distance of observations from the predictor, while quantile estimation only knows whether an observation is below or above the predictor. It would be awkward to measure extreme risk based only on the frequency of tail losses and not on their values. An interesting asymptotic connection between XMES(τ) and QMES(τ) is provided below in Proposition 5. It is also the goal of the next section to establish estimators of the tail expectile-based MES and to unravel their asymptotic behavior. The asymptotic normality is derived for a large class of bivariate distributions of (X,Y), which makes statistical inference for XMES(τ) feasible.

5.2 Tail dependence model

Suppose the random vector (X,Y) has a continuous bivariate distribution function $F_{(X,Y)}$ and denote by F_X and F_Y the marginal distribution functions of X and Y, assumed to be increasing in what follows. Given that our goal is to estimate XMES(τ) at an extreme level τ, we adopt the same conditions as Cai *et al.* (2015) on the right-hand tail of X and on the right-hand upper tail dependence of (X, Y). Here, the right-hand upper tail dependence between X and Y is described by the following joint convergence condition:

$\boldsymbol{\mathcal{JC}(R)}$ For all $(x, y) \in [0, \infty]^2$ such that at least x or y is finite, the limit

$$\lim_{t\to\infty} t\mathbb{P}(\overline{F}_X(X) \leqslant x/t, \overline{F}_Y(Y) \leqslant y/t) := R(x, y)$$

exists, with $\overline{F}_X = 1 - F_X$ and $\overline{F}_Y = 1 - F_Y$.

The limit function R completely determines the so-called tail dependence function ℓ [Drees and Huang (1998)] via the identity $\ell(x, y) = x + y - R(x, y)$ for all $x, y \geqslant 0$ [see also Beirlant *et al.* (2004), Section 8.2]. Regarding the marginal distributions, we assume that X and Y are heavy-tailed with respective tail indices $\gamma_X, \gamma_Y > 0$, or equivalently, for all $z > 0$,

$$\frac{U_X(tz)}{U_X(t)} \to z^{\gamma_X} \quad \text{and} \quad \frac{U_Y(tz)}{U_Y(t)} \to z^{\gamma_Y} \quad \text{as} \quad t \to \infty,$$

with U_X and U_Y being, respectively, the left-continuous inverse functions of $1/\overline{F}_X$ and $1/\overline{F}_Y$. Compared with the quantile-based MES framework in Cai *et al.* (2015), we need the extra condition of heavy-tailedness of Y which is quite natural in the financial setting. Under these regularity conditions, we get the following asymptotic approximations for XMES(τ).

Proposition 5. *Suppose that condition $\mathcal{JC}(R)$ holds and that X and Y are heavy-tailed with respective indices $\gamma_X, \gamma_Y \in (0, 1)$. Then*

$$\lim_{\tau\uparrow 1} \frac{XMES(\tau)}{U_X(1/\overline{F}_Y(\xi_{Y,\tau}))} = \int_0^\infty R(x^{-1/\gamma_X}, 1)dx, \tag{19}$$

$$\lim_{\tau\uparrow 1} \frac{XMES(\tau)}{QMES(\tau)} = \left(\gamma_Y^{-1} - 1\right)^{-\gamma_X}. \tag{20}$$

The first convergence result indicates that XMES(τ) is asymptotically equivalent to the small exceedance probability $U_X(1/\overline{F}_Y(\xi_{Y,\tau}))$ up to a multiplicative constant. Since as usual in the financial setting $0 < \gamma_X, \gamma_Y < 1/2$, the second result shows that XMES(τ) is less extreme than QMES(τ) as $\tau \to 1$. This is visualised in Figure 2 in the case of a standard bivariate Student t_ν-distribution on $(0, \infty)^2$ with density

$$f_\nu(x, y) = \frac{2}{\pi}\left(1 + \frac{x^2 + y^2}{\nu}\right)^{-(\nu+2)/2}, \quad x, y > 0, \tag{21}$$

where $\nu = 3, 5$, respectively from left to right. It can be seen that QMES(τ) becomes overall much more extreme than XMES(τ) as τ approaches 1.

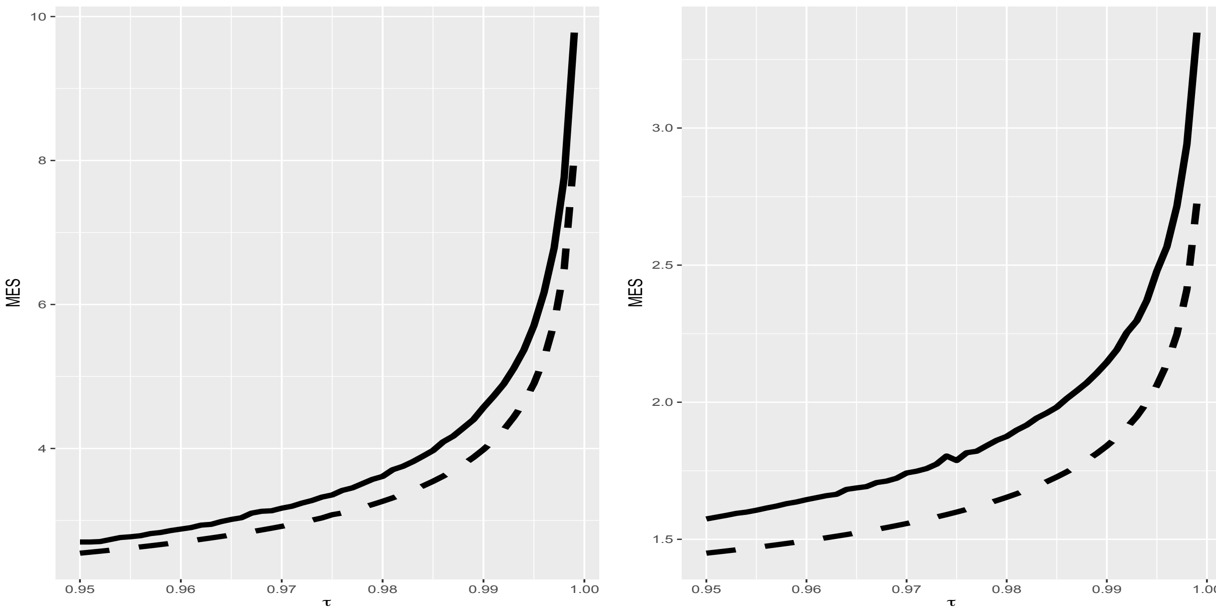

Figure 2: *QMES*(τ) *in solid line and XMES*(τ) *in dashed line, as functions of* $\tau \in [0.95, 1)$. *Case of Student* t_ν*-distribution on* $(0, \infty)^2$. *From left to right,* $\nu = 3, 5$.

5.3 Estimation and results

The asymptotic equivalences in Proposition 5 are of particular interest when it comes to proposing estimators for tail expectile-based MES. Two approaches will be distinguished. We consider first asymmetric least squares estimation by making use of the asymptotic equivalence (19). Subsequently we shall deal with a nonparametric estimator derived from the asymptotic connection (20) with the tail quantile-based MES.

5.3.1 Asymmetric least squares estimation

On the basis of the limit (19) and then of the heavy-tailedness assumption on X, we have for $\tau < \tau' < 1$ that, as $\tau \to 1$,

$$\text{XMES}(\tau') \approx \frac{U_X(1/\overline{F}_Y(\xi_{Y,\tau'}))}{U_X(1/\overline{F}_Y(\xi_{Y,\tau}))}\text{XMES}(\tau) \approx \left(\frac{\overline{F}_Y(\xi_{Y,\tau})}{\overline{F}_Y(\xi_{Y,\tau'})}\right)^{\gamma_X} \text{XMES}(\tau).$$

It follows then from Proposition 1 that

$$\text{XMES}(\tau') \approx \left(\frac{1-\tau'}{1-\tau}\right)^{-\gamma_X} \text{XMES}(\tau). \tag{22}$$

Hence, to estimate $\text{XMES}(\tau')$ at an arbitrary extreme level $\tau' = \tau_n'$, we first consider the estimation of $\text{XMES}(\tau)$ at an intermediate level $\tau = \tau_n$, and then we use the extrapolation technique of Weissman (1978). For estimating $\text{XMES}(\tau_n) = \mathbb{E}\{X|Y > \xi_{Y,\tau_n}\}$ at an intermediate level $\tau_n \to 1$ such that $n(1-\tau_n) \to \infty$, as $n \to \infty$, we use the empirical version

$$\widetilde{\mathrm{XMES}}(\tau_n) := \frac{\sum_{i=1}^n X_i \mathbb{1}\{X_i > 0, Y_i > \widetilde{\xi}_{Y,\tau_n}\}}{\sum_{i=1}^n \mathbb{1}\{Y_i > \widetilde{\xi}_{Y,\tau_n}\}},$$

where $\widetilde{\xi}_{Y,\tau_n}$ is the LAWS estimator of ξ_{Y,τ_n}. As a matter of fact, in actuarial settings, we typically have a positive loss variable X, and hence $\mathbb{1}\{X_i > 0\} = 1$. When considering a real-valued profit-loss variable X, the MES is mainly determined by high, and hence positive, values of X as shown in Cai *et al.* (2015).

We shall show under general conditions that the estimator $\widetilde{\mathrm{XMES}}(\tau_n)$ is $\sqrt{n(1-\tau_n)}$-relatively consistent. By plugging this estimator into approximation (22) together with a $\sqrt{n(1-\tau_n)}$-consistent estimator $\widehat{\gamma}_X$ of γ_X, we obtain the following estimator of $\mathrm{XMES}(\tau_n')$:

$$\widetilde{\mathrm{XMES}}^\star(\tau_n') \equiv \widetilde{\mathrm{XMES}}^\star(\tau_n';\tau_n) := \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_X} \widetilde{\mathrm{XMES}}(\tau_n). \tag{23}$$

To determine the limit distribution of this estimator, we need to quantify the rate of convergence in condition $\mathcal{JC}(R)$ as follows:

$\boldsymbol{\mathcal{JC}_2(R,\beta,\kappa)}$ Condition $\mathcal{JC}(R)$ holds and there exist $\beta > \gamma_X$ and $\kappa < 0$ such that

$$\sup_{\substack{x\in(0,\infty)\\ y\in[1/2,2]}} \left| \frac{t\mathbb{P}(\overline{F}_X(X) \leqslant x/t, \overline{F}_Y(Y) \leqslant y/t) - R(x,y)}{\min(x^\beta, 1)} \right| = \mathrm{O}(t^\kappa) \quad \text{as} \quad t \to \infty.$$

This is exactly condition (a) in Cai *et al.* (2015) under which an extrapolated estimator of $\mathrm{QMES}(\tau_n')$ converges to a normal distribution. See also condition (7.2.8) in de Haan and Ferreira (2006). We also need to assume that the tail quantile function U_X (resp. U_Y) satisfies the second-order condition $\mathcal{C}_2(\gamma_X, \rho_X, A_X)$ (resp. $\mathcal{C}_2(\gamma_Y, \rho_Y, A_Y)$). The following generic theorem gives the asymptotic distribution of $\widetilde{\mathrm{XMES}}^\star(\tau_n')$. The asymptotic normality follows by using for example the Hill estimator $\widehat{\gamma}_X$ of the tail index γ_X.

Theorem 4. *Suppose that condition $\mathcal{JC}_2(R,\beta,\kappa)$ holds, that there is $\delta > 0$ such that $\mathbb{E}|Y_-|^{2+\delta} < \infty$, and that U_X and U_Y satisfy conditions $\mathcal{C}_2(\gamma_X,\rho_X,A_X)$ and $\mathcal{C}_2(\gamma_Y,\rho_Y,A_Y)$ with $\gamma_X, \gamma_Y \in (0,1/2)$ and $\rho_X < 0$. Assume further that*

(i) τ_n, $\tau_n' \uparrow 1$, *with* $n(1-\tau_n) \to \infty$, $n(1-\tau_n') \to c < \infty$ *and* $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau_n')] \to \infty$ *as* $n \to \infty$;

(ii) $1-\tau_n = \mathrm{O}(n^{\alpha-1})$ *for some* $\alpha < \min\left(\dfrac{-2\kappa}{-2\kappa+1}, \dfrac{2\gamma_X\rho_X}{2\gamma_X\rho_X+\rho_X-1}\right)$;

(iii) *The bias conditions* $\sqrt{n(1-\tau_n)}q_{Y,\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}A_X((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$ *and* $\sqrt{n(1-\tau_n)}A_Y((1-\tau_n)^{-1}) \to \lambda_3 \in \mathbb{R}$ *hold;*

(iv) $\sqrt{n(1-\tau_n)}(\widehat{\gamma}_X - \gamma_X) \xrightarrow{d} \Gamma.$

Then, if $X > 0$ almost surely, we have that

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widetilde{XMES}^{\star}(\tau_n')}{XMES(\tau_n')} - 1\right) \xrightarrow{d} \Gamma.$$

This convergence remains still valid if $X \in \mathbb{R}$ provided

(v) $\mathbb{E}|X_-|^{1/\gamma_X} < \infty;$ (24)

(vi) $n(1-\tau_n) = \mathrm{o}\left((1-\tau_n')^{2\kappa(1-\gamma_X)}\right)$ *as* $n \to \infty.$ (25)

Let us point out here that condition (ii), which also appears in Theorem 1 of Cai *et al.* (2015), is a strengthening of the condition $1 - \tau_n = \mathrm{o}(1)$. It essentially allows to control additional bias terms that appear in conditions $\mathcal{JC}_2(R,\beta,\kappa)$ and $\mathcal{C}_2(\gamma_X, \rho_X, A_X)$. Condition (vi), which is also utilized in Cai *et al.* (2015), is another bias condition that makes it possible to control the bias coming from the left tail of X.

5.3.2 Estimation based on tail QMES

On the basis of the limit (20), we consider the alternative estimator

$$\widehat{\mathrm{XMES}}^{\star}(\tau_n') := \left(\widehat{\gamma}_Y^{-1} - 1\right)^{-\widehat{\gamma}_X} \widehat{\mathrm{QMES}}^{\star}(\tau_n'), \tag{26}$$

where $\widehat{\gamma}_X$, $\widehat{\gamma}_Y$ and $\widehat{\mathrm{QMES}}^{\star}(\tau_n')$ are suitable estimators of γ_X, γ_Y and $\mathrm{QMES}(\tau_n')$, respectively. Here, we use the Weissman-type device

$$\widehat{\mathrm{QMES}}^{\star}(\tau_n') = \left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_X} \widehat{\mathrm{QMES}}(\tau_n) \tag{27}$$

of Cai *et al.* (2015) to estimate $\mathrm{QMES}(\tau_n')$, where

$$\widehat{\mathrm{QMES}}(\tau_n) = \frac{1}{\lfloor n(1-\tau_n)\rfloor}\sum_{i=1}^{n} X_i \mathbb{1}\{X_i > 0, Y_i > \widehat{q}_{Y,\tau_n}\},$$

with $\widehat{q}_{Y,\tau_n} := Y_{n-\lfloor n(1-\tau_n)\rfloor,n}$ being an intermediate quantile-VaR. As a matter of fact, Cai *et al.* (2015) have suggested the use of two intermediate sequences in $\widehat{\gamma}_X$ and $\widehat{\mathrm{QMES}}(\tau_n)$ to be chosen in two steps in practice. To ease the presentation, we use the same intermediate sequence τ_n in both $\widehat{\gamma}_X$ and $\widehat{\mathrm{QMES}}(\tau_n)$. Next, we derive the asymptotic distribution of the new estimator $\widehat{\mathrm{XMES}}^{\star}(\tau_n')$.

Theorem 5. *Suppose that condition $\mathcal{JC}_2(R,\beta,\kappa)$ holds, and U_X and U_Y satisfy conditions $\mathcal{C}_2(\gamma_X,\rho_X,A_X)$ and $\mathcal{C}_2(\gamma_Y,\rho_Y,A_Y)$ with $\gamma_X \in (0,1/2)$ and $\rho_X < 0$. Assume further that* *(i)* τ_n*,* $\tau_n' \uparrow 1$*, with* $n(1-\tau_n) \to \infty$*,* $n(1-\tau_n') \to c < \infty$ *and* $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\tau_n')] \to \infty$ *as* $n \to \infty$*;*

(ii) $1-\tau_n = \mathrm{O}(n^{\alpha-1})$ *for some* $\alpha < \min\left(\dfrac{-2\kappa}{-2\kappa+1}, \dfrac{2\gamma_X\rho_X}{2\gamma_X\rho_X+\rho_X-1}\right)$*;*

(iii) The bias conditions $\sqrt{n(1-\tau_n)}q_{Y,\tau_n}^{-1} \to \lambda \in \mathbb{R}$ *and* $\sqrt{n(1-\tau_n)}A_X((1-\tau_n)^{-1}) \to 0$ *hold;*

(iv) $\sqrt{n(1-\tau_n)}(\widehat{\gamma}_X - \gamma_X) \xrightarrow{d} \Gamma$ *and* $\sqrt{n(1-\tau_n)}(\widehat{\gamma}_Y - \gamma_Y) = \mathrm{O}_{\mathbb{P}}(1)$*.*

Then, if $X > 0$ *almost surely, we have that*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau_n')]}\left(\frac{\widehat{XMES}^{\star}(\tau_n')}{XMES(\tau_n')} - 1\right) \xrightarrow{d} \Gamma.$$

This convergence remains still valid if $X \in \mathbb{R}$ *provided that* (24) *and* (25) *hold.*

6 Extreme expectile level selection

An important question that remains to be addressed is the choice of the extreme expectile level τ_n' in the three instruments of risk protection $\xi_{\tau_n'}$, $\mathrm{XES}(\tau_n')$ and $\mathrm{XMES}(\tau_n')$.

In the case of quantile-based risk measures q_{α_n}, $\mathrm{QES}(\alpha_n)$ and $\mathrm{QMES}(\alpha_n)$, it is customary to choose tail probabilities $\alpha_n \to 1$ with $n(1-\alpha_n) \to c$, a finite constant, as the sample size $n \to \infty$, to allow for more 'prudent' risk management. In response to the many turbulent episodes that have been experienced by financial markets during the last few decades, academics are nowadays more interested in once-in-a-decade or twice-per-decade events (see, *e.g.*, Brownlees and Engle (2016) and Cai *et al.* (2015)). In the case of expectiles, we propose to select τ_n' so that each expectile-based risk measure has the same intuitive interpretation as its quantile-based analogue. This translates into choosing τ_n' such that $\xi_{\tau_n'} \equiv q_{\alpha_n}$ for a given relative frequency α_n. Bellini and Di Bernardino (2015) have already suggested to pick out τ_n' which satisfies $\xi_{\tau_n'} = q_{\alpha_n}$, but for a normally distributed Y. Here, we wish to extend this elegant device to a general random variable Y without any a priori specification.

Thanks to the connection (2), it is immediate from $\xi_{\tau_n'} \equiv q_{\alpha_n}$ that $\tau_n'(\alpha_n) := \tau_n'$ satisfies

$$1 - \tau_n'(\alpha_n) = \frac{\mathbb{E}\left\{|Y - q_{\alpha_n}|\, \mathbb{1}(Y > q_{\alpha_n})\right\}}{\mathbb{E}\,|Y - q_{\alpha_n}|}.$$

As a matter of fact, under the model assumption of Pareto-type tails, it turns out that the expectile level $\tau_n'(\alpha_n)$ depends asymptotically only on the quantile level α_n and not on the quantile q_{α_n} itself.

Proposition 6. *Suppose F_Y satisfies* (3) *with* $0 < \gamma < 1$. *Then*

$$1 - \tau_n'(\alpha_n) \sim (1 - \alpha_n)\,\frac{\gamma}{1-\gamma}, \quad n \to \infty.$$

Hence, by substituting the estimated value

$$\widehat{\tau}_n'(\alpha_n) = 1 - (1 - \alpha_n)\,\frac{\widehat{\gamma}}{1 - \widehat{\gamma}}$$

in place of τ_n', both extreme expectile estimators $\widehat{\xi}^{\star}_{\tau_n'}$ in (11) and $\widetilde{\xi}^{\star}_{\tau_n'}$ in (12) estimate the same Value at Risk $\xi_{\tau_n'(\alpha_n)} \equiv q_{\alpha_n}$ as the Weissman quantile estimator $\widehat{q}^{\star}_{\alpha_n}$ in (10). It is easily seen that the latter estimator is actually identical to the indirect expectile estimator $\widehat{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)}$. Indeed, we have in view of (10) and (11) that

$$\begin{aligned}
\widehat{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)} = \left(\widehat{\gamma}^{-1} - 1\right)^{-\widehat{\gamma}} \widehat{q}^{\star}_{\widehat{\tau}_n'(\alpha_n)} &= \left(\widehat{\gamma}^{-1} - 1\right)^{-\widehat{\gamma}} \left(\frac{1 - \widehat{\tau}_n'(\alpha_n)}{1 - \tau_n}\right)^{-\widehat{\gamma}} \widehat{q}_{\tau_n} \\
&= \left(\frac{1 - \widehat{\gamma}}{\widehat{\gamma}}\right)^{-\widehat{\gamma}} \left(\frac{(1 - \alpha_n)\frac{\widehat{\gamma}}{1-\widehat{\gamma}}}{1 - \tau_n}\right)^{-\widehat{\gamma}} \widehat{q}_{\tau_n} \\
&= \left(\frac{1 - \alpha_n}{1 - \tau_n}\right)^{-\widehat{\gamma}} \widehat{q}_{\tau_n} \\
&= \widehat{q}^{\star}_{\alpha_n}.
\end{aligned}$$

This quantile-based estimator $\widehat{q}^{\star}_{\alpha_n} \equiv \widehat{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)}$ may be criticized for being optimistic (or liberal) because it relies on a single order statistic $\widehat{q}_{\tau_n} = Y_{n-\lfloor n(1-\tau_n)\rfloor,n}$, and hence may not respond properly to the very extreme losses. By contrast, the direct expectile-based estimator $\widetilde{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)}$ relies on the asymmetric least squares estimator $\widetilde{\xi}_{\tau_n}$, and hence bears much better the burden of representing a sensitive risk measure to the magnitude of infrequent catastrophic losses. The next result shows that the asymptotic behavior of the original extrapolated estimators $\widehat{\xi}^{\star}_{\tau_n'}$ and $\widetilde{\xi}^{\star}_{\tau_n'}$, established in Corollaries 3 and 4, remains still valid for the resulting composite estimators $\widehat{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)}$ and $\widetilde{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)}$, under the same technical conditions.

Theorem 6. *(i) Assume that F_Y is strictly increasing, that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds with $0 < \gamma < 1$ and $\rho < 0$, and that τ_n, $\alpha_n \uparrow 1$ with $n(1 - \tau_n) \to \infty$ and $n(1 - \alpha_n) \to c < \infty$. Assume further that*

$$\sqrt{n(1 - \tau_n)} \left(\widehat{\gamma} - \gamma, \frac{\widehat{q}_{\tau_n}}{q_{\tau_n}} - 1\right) \xrightarrow{d} (\Gamma, \Theta).$$

If $\sqrt{n(1 - \tau_n)}q^{-1}_{\tau_n} \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1 - \tau_n)}A((1 - \tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$ and $\sqrt{n(1 - \tau_n)}/\log[(1 - \tau_n)/(1 - \alpha_n)] \to \infty$, then

$$\frac{\sqrt{n(1 - \tau_n)}}{\log[(1 - \tau_n)/(1 - \alpha_n)]} \left(\frac{\widehat{\xi}^{\star}_{\widehat{\tau}_n'(\alpha_n)}}{q_{\alpha_n}} - 1\right) \xrightarrow{d} \Gamma.$$

(ii) Assume that F_Y is strictly increasing, that there is $\delta > 0$ such that $\mathbb{E}|Y_-|^{2+\delta} < \infty$, that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds with $0 < \gamma < 1/2$ and $\rho < 0$, and that τ_n, $\alpha_n \uparrow 1$ with $n(1-\tau_n) \to \infty$ and $n(1-\alpha_n) \to c < \infty$. If in addition

$$\sqrt{n(1-\tau_n)}(\widehat{\gamma} - \gamma) \xrightarrow{d} \Gamma$$

and $\sqrt{n(1-\tau_n)}q_{\tau_n}^{-1} \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_2 \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/\log[(1-\tau_n)/(1-\alpha_n)] \to \infty$, then

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\alpha_n)]}\left(\frac{\widetilde{\xi}^{\star}_{\widehat{\tau}'_n(\alpha_n)}}{q_{\alpha_n}} - 1\right) \xrightarrow{d} \Gamma.$$

Likewise, $\widehat{\text{XES}}^{\star}(\tau'_n)$ and $\widetilde{\text{XES}}^{\star}(\tau'_n)$ in (17) as well as $\widehat{\text{XES}}^{\dagger}(\tau'_n)$ and $\widetilde{\text{XES}}^{\dagger}(\tau'_n)$ in (18), with $\tau'_n = \widehat{\tau}'_n(\alpha_n)$, estimate the same expected shortfall $\text{XES}(\tau'_n(\alpha_n)) \equiv \text{QES}(\alpha_n)$ as the quantile-based estimator $\widehat{\text{QES}}^{\star}(\alpha_n)$ described in (28). Note also that $\widehat{\text{XES}}^{\dagger}(\widehat{\tau}'_n(\alpha_n))$ coincides with $\widehat{\text{QES}}^{\star}(\alpha_n)$. Moreover, our numerical illustrations indicate that $\widehat{\text{QES}}^{\star}(\alpha_n)$ points towards similar estimates as $\widehat{\text{XES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$, but the direct expectile-based estimators $\widetilde{\text{XES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ and $\widetilde{\text{XES}}^{\dagger}(\widehat{\tau}'_n(\alpha_n))$ tend to be more conservative. The next result can be proved by making use of the proof of the previous Theorem just as Corollaries 5 and 6 follow from the proofs of Corollaries 3 and 4.

Theorem 7. *(i) Under the conditions of Theorem 6 (i),*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\alpha_n)]}\left(\frac{\widehat{XES}^{\star}(\widehat{\tau}'_n(\alpha_n))}{QES(\alpha_n)} - 1\right) \xrightarrow{d} \Gamma.$$

(ii) Under the conditions of Theorem 6 (ii),

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\alpha_n)]}\left(\frac{\widetilde{XES}^{\star}(\widehat{\tau}'_n(\alpha_n))}{QES(\alpha_n)} - 1\right) \xrightarrow{d} \Gamma.$$

Let us now turn to $\widetilde{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ in (23) and $\widehat{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ in (26) that estimate the same marginal expected shortfall $\text{XMES}(\tau'_n(\alpha_n)) \equiv \text{QMES}(\alpha_n)$ as Cai *et al.* (2015)'s estimator $\widehat{\text{QMES}}^{\star}(\alpha_n)$ defined in (27). As a matter of fact, $\widehat{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ is nothing but $\widehat{\text{QMES}}^{\star}(\alpha_n)$.

Theorem 8. *(i) Suppose the conditions of Theorem 4 hold with α_n in place of τ'_n. Then*

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\alpha_n)]}\left(\frac{\widetilde{XMES}^{\star}(\widehat{\tau}'_n(\alpha_n))}{QMES(\alpha_n)} - 1\right) \xrightarrow{d} \Gamma.$$

(ii) Suppose the conditions of Theorem 5 hold with α_n in place of τ'_n. Then

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\alpha_n)]}\left(\frac{\widehat{XMES}^{\star}(\widehat{\tau}'_n(\alpha_n))}{QMES(\alpha_n)} - 1\right) \xrightarrow{d} \Gamma.$$

7 Simulation study

The aim of this section is to highlight some of the theoretical findings with numerical simulations. We will briefly touch on the presented tail XVaR and XES estimators in section 7.1 and tail XMES estimators in section 7.2. Both sections provide Monte-Carlo evidence that the direct estimation method is more efficient relative to the indirect method in the case of real-valued profit-loss variables, whereas the rival indirect method tends to be the winner in the case of non-negative loss distributions. The latter method seems to be also superior in the case of extremely heavy tails.

7.1 Expectile-based VaR and ES

To evaluate finite-sample performance of the extreme expectile estimators $\widetilde{\xi}^{\star}_{\tau_n'} \equiv \widetilde{\xi}^{\star}_{\tau_n'}(\tau_n)$ and $\widehat{\xi}^{\star}_{\tau_n'} \equiv \widehat{\xi}^{\star}_{\tau_n'}(\tau_n)$, we have considered simulated samples from the Student t_ν-distribution ($\nu = 3, 5, 7, 9$), which corresponds to real-valued profit-loss variables, and from the marginal of the bivariate Student t_ν-distribution described in (21), which corresponds to non-negative loss variables. We shall refer to this marginal distribution on $(0, \infty)$ as 'positive Student t_ν-distribution'. We used in all our simulations the Hill estimator of γ, the extreme level $\tau_n' = 0.995$ for $n = 100$ and $\tau_n' = 0.9994$ for $n = 1000$, and the intermediate levels $\tau_n = 1 - \frac{k}{n}$, where the integer k can actually be viewed as the effective sample size for tail extrapolation. We only present here the results for $n = 1000$ and $\nu \in \{3, 5\}$, a full comparison including additional results for optimal k is given in Supplement A.1.

In the case of Student t-distributions, Figure 3 gives the root Mean-Squared Error (MSE) in top panels and bias estimates in bottom panels, computed over $10,000$ replications for samples of size 1000. Each figure displays the evolution of the obtained Monte-Carlo results, for the two normalized estimators $\widetilde{\xi}^{\star}_{\tau_n'}(k)/\xi_{\tau_n'}$ and $\widehat{\xi}^{\star}_{\tau_n'}(k)/\xi_{\tau_n'}$, as functions of the sample fraction k. Our tentative conclusion is that the accuracy of the direct estimator $\widetilde{\xi}^{\star}_{\tau_n'}$ is quite respectable relative to the indirect estimator $\widehat{\xi}^{\star}_{\tau_n'}$. Our experience with other simulated data indicates, however, that the direct estimator is no longer the winner in the case of extremely heavy-tailed distributions such as, for instance, Student t_ν-distributions with $1 < \nu \leqslant 2$.

The resulting Monte-Carlo estimates in the case of positive Student distributions, displayed in Figure 4, indicate that the indirect estimator $\widehat{\xi}^{\star}_{\tau_n'}$ is superior to the direct estimator $\widetilde{\xi}^{\star}_{\tau_n'}$: the use of the Pareto distribution $F_Y(y) = 1 - y^{-1/\gamma}$, $y > 1$, and the Fréchet distribution $F_Y(y) = e^{-y^{-1/\gamma}}$, $y > 0$, leads to the same conclusion. It may also be seen in both Student and positive Student scenarios that most of the error is due to variance, the squared bias being much smaller in all cases. This may be explained by the sensitivity of high expectiles to the magnitude of heavy tails, since they are based on "squared" error loss minimization. It is interesting that in almost all cases the bias was positive.

Another way of validating the presented estimation procedures for $\xi_{\tau'_n}$ on historical data is by using the elicitability property of expectiles as pointed out in Section 1. Following the ideas of Gneiting (2011), the competing estimates $\widehat{\xi}^{\star}_{\tau'_n}$ and $\widetilde{\xi}^{\star}_{\tau'_n}$ can be compared from a forecasting perspective by means of their realized losses. A more comprehensive description of this comparison including Monte Carlo verification and validation is given in Supplement A.2, where the resulting average values of the realized losses seem to favor $\widetilde{\xi}^{\star}_{\tau'_n}$ over $\widehat{\xi}^{\star}_{\tau'_n}$ in the case of Student t-distributions, while they tend to prefer $\widehat{\xi}^{\star}_{\tau'_n}$ over $\widetilde{\xi}^{\star}_{\tau'_n}$ in the case of positive Student t-distributions. Note that one can also compare the two forecasters $\widehat{\xi}^{\star}_{\tau'_n}$ and $\widetilde{\xi}^{\star}_{\tau'_n}$ by applying the modern and promising tool of 'Murphy diagram', recently developed by Ehm *et al.* (2016).

We also investigate the normality of the estimators $\widehat{\xi}^{\star}_{\tau'_n}$ and $\widetilde{\xi}^{\star}_{\tau'_n}$ in Supplement A.3, where the Q–Q-plots indicate that the limit Theorem 3 and its Corollaries 3 and 4 provide adequate approximations for finite sample sizes.

Other simulation experiments have been undertaken to assess the finite-sample performance of the expectile-based ES estimators $\widehat{\text{XES}}^{\star}(\tau'_n)$, $\widetilde{\text{XES}}^{\star}(\tau'_n)$, $\widehat{\text{XES}}^{\dagger}(\tau'_n)$ and $\widetilde{\text{XES}}^{\dagger}(\tau'_n)$. The experiments all employed the same families of Student and positive Student t-distributions as before. The lessons were similar to those from the expectile-based VaR setting, hence the results are not reported here. It may also be noticed that the Monte-Carlo estimates corresponding to $\widehat{\text{XES}}^{\star}(\tau'_n)$ and $\widehat{\text{XES}}^{\dagger}(\tau'_n)$ [respectively, $\widetilde{\text{XES}}^{\star}(\tau'_n)$ and $\widetilde{\text{XES}}^{\dagger}(\tau'_n)$] are very similar.

7.2 Expectile-based MES

Here, we compare the composite estimators $\widetilde{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ and $\widehat{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ that estimate the same MES, $\text{XMES}(\tau'_n(\alpha_n)) \equiv \text{QMES}(\alpha_n)$, as the Cai *et al.* (2015) estimator $\widehat{\text{QMES}}^{\star}(\alpha_n)$. The latter is actually identical to the indirect estimator $\widehat{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$. All the experiments have sample size $n = 1000$ and extreme level $\alpha_n = 0.9994$.

To investigate the finite sample performance of the two rival estimators $\widetilde{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ and $\widehat{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$, the simulation experiments first employ the Student t_ν-distribution on $(0,\infty)^2$ with density $f_\nu(x,y)$ described in (21). It can be shown that this distribution satisfies the conditions $\mathcal{JC}_2(R,\beta,\kappa)$ and $\mathcal{C}_2(\gamma_X,\rho_X,A_X)$ of Theorems 4 and 5 (see Cai *et al.* (2015) for the case $\nu = 3$). Other motivating examples of distributions that satisfy these conditions can also be found in section 3 of Cai *et al.* (2015). All the experiments have $\nu \in \{3,5,7,9\}$. As they point towards the same conclusions, we only present the results for $\nu = 3,5$. For the choice of the intermediate level τ_n, we used the same considerations as in Section 7.1.

In Figure 5 we present the root-MSE (top panels) and bias estimates (bottom panels) computed over 10,000 simulated samples. Each picture displays the evolution of the obtained Monte-Carlo results, for the two normalized estimators $\widetilde{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))/\text{XMES}(\tau'_n(\alpha_n))$ and $\widehat{\text{XMES}}^{\star}(\widehat{\tau}'_n(\alpha_n))/\text{XMES}(\tau'_n(\alpha_n))$, as functions of the sample fraction k. We observe that the latter indirect estimator is clearly the winner in all cases in terms of both root-MSE and bias. As can also be seen in Supplement A.3, the limit Theorems 4 and 5 provide adequate approximations for finite sample sizes, with a slight advantage for $\widehat{\text{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$.

To illustrate the case of real-valued profit-loss random variables, we consider a transformed Student t_ν-distribution on the whole of the plane $\mathbb{R}^2$ defined as

$$(X,Y) = \left(Z_1^{\nu/4} 1\!\!1(Z_1 \geqslant 0) - (-Z_1)^{\nu/8} 1\!\!1(Z_1 < 0), Z_2 \right),$$

where (Z_1, Z_2) denotes a standard Student t_ν-distribution on $\mathbb{R}^2$ with density

$$\frac{1}{2\pi} \left(1 + \left(x^2 + y^2\right)/\nu\right)^{-(\nu+2)/2}, \quad x, y \in \mathbb{R}.$$

The resulting Monte-Carlo estimates for $\nu \in \{3, 5\}$, displayed in Figure 6, indicate that $\widetilde{\text{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$ is more efficient relative to $\widehat{\text{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$. This superiority of the direct estimator is, however, no more longer valid in the case of extremely heavy tails such as, for instance, $\nu = 2$ and the transformed Cauchy distribution considered in Cai *et al.* (2015).

8 Applications

In this section, we apply our estimation methods to first estimate the tail VaR and ES for the Society of Actuaries (SOA) Group Medical Insurance Large Claims, and then to estimate the tail MES for three large investment banks in the USA.

8.1 VaR and ES for medical insurance data

The SOA Group Medical Insurance Large Claims Database records all the claim amounts exceeding 25,000 USD over the period 1991-92. As in Beirlant *et al.* (2004), we only deal here with the 75,789 claims for 1991. The histogram and scatterplot shown in Figure 7 (a) give evidence of an important right-skewness. Insurance companies are then interested in estimating the worst tail value of the corresponding loss severity distribution. One way of measuring this value at risk is by considering the Weissman quantile estimate $\widehat{q}^{\star}_{\alpha_n} = Y_{n-k,n}\left(\frac{k}{np_n}\right)^{\widehat{\gamma}_H}$ as described in (10), where $\widehat{\gamma}_H$ is the Hill estimator defined in (6), with $\alpha_n = 1 - p_n$ and $\tau_n = 1 - \frac{k}{n}$. According to the earlier study of Beirlant *et al.* (2004, p.123), insurers typically are interested in $p_n = \frac{1}{100{,}000} \approx \frac{1}{n}$ for these medical insurance data, that is, in an estimate of the claim amount that will be exceeded (on average) only once in 100,000 cases. Similar recent studies in the context of the backtesting problem, which is crucial in the current Basel III regulatory framework (see Basel Committee, 2011), are Chavez-Demoulin *et al.* (2014) and Gong *et al.* (2015), who estimate quantiles exceeded on average once every 100 cases with sample sizes of the order of hundreds. Figure 7 (b) shows the quantile-VaR estimates $\widehat{q}^{\star}_{\alpha_n}$ against the sample fraction k (solid line). A commonly used heuristic approach for selecting a pointwise estimate is to pick out a value of k corresponding to the first stable part of the plot [see, *e.g.*, Section 3 in de Haan and Ferreira (2006)]. Here, a stable region appears for k from 150 up to 500, leading to an estimate between 3.73 and 4.12 million. This estimate does not succeed in exceeding the sample maximum $Y_{n,n} = 4,518,420$ (indicated by the horizontal line), which is consistent with the earlier analysis of Beirlant *et al.* (2004, p.125 and p.159).

The alternative expectile-based estimator $\widetilde{\xi}^{\star}_{\widehat{\tau}'_n(\alpha_n)}$ introduced in Section 6, which estimates the same VaR $q_{\alpha_n} \equiv \xi_{\tau'_n(\alpha_n)}$ as the quantile-based estimator $\widehat{q}^{\star}_{\alpha_n} \equiv \widehat{\xi}^{\star}_{\widehat{\tau}'_n(\alpha_n)}$, is also graphed in Figure 7 (b) in dashed line. As an asymmetric-least-squares estimator, it is more affected by the infrequent great claim amounts visualized in the top figure. Its plot indicates a more conservative risk measure between 3.92 and 4.33 million, over the stable region $k \in [150, 500]$. Yet, this measure is less severe than the maximal recorded claim amount $Y_{n,n}$.

An alternative option for measuring risk, which is more capable of extrapolating outside the range of the available observations, is by using the estimated quantile-ES

$$\widehat{\mathrm{QES}}^{\star}(\alpha_n) = \frac{1}{k}\sum_{i=1}^{n} Y_i \mathbb{1}(Y_i > Y_{n-k,n}) \cdot \left(\frac{k}{np_n}\right)^{\widehat{\gamma}_H} \tag{28}$$

[see El Methni *et al.* (2014)]. Its graph shown in Figure 7 (c) in dashed-dotted line indicates a stable region for $k \in [150, 500]$ with an averaged estimate of around 6.13 million, which is successfully extrapolated beyond the data. The graph of $\widehat{\mathrm{QES}}^{\star}(\alpha_n)$ is very close to the plot in solid line of the indirect expectile-based estimator $\widehat{\mathrm{XES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ which estimates the same risk measure $\mathrm{XES}(\tau'_n(\alpha_n)) \equiv \mathrm{QES}(\alpha_n)$. In contrast to $\widehat{\mathrm{XES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$, which indicates an averaged risk estimate of around 6.14 million over the stable region $k \in [150, 500]$, both direct expectile-based estimators $\widetilde{\mathrm{XES}}^{\star}(\widehat{\tau}'_n(\alpha_n))$ in dashed line and $\widetilde{\mathrm{XES}}^{\dagger}(\widehat{\tau}'_n(\alpha_n))$ in dotted line are clearly more pessimistic as they rely on the asymmetric-least-squares estimator $\widetilde{\xi}^{\star}_{\widehat{\tau}'_n(\alpha_n)}$. The averaged values of these pessimistic ES estimates over $k \in [150, 500]$ are, respectively, around 6.5 and 6.48 million. They exceed the traditional quantile-ES estimate by 0.37 million.

That eternal maxim of the cautious aunt and misanthropic uncle, *"expect the worst, and you won't be disappointed"* [Bassett *et al.* (2004)] might thus be transformed into a concrete calculus via the asymmetric least squares-based VaR and ES estimates.

8.2 MES of three large US financial institutions

We consider the same investment banks as in the studies of Brownlees and Engle (2016) and Cai *et al.* (2015), namely Goldman Sachs, Morgan Stanley and T. Rowe Price. For the three banks, the dataset consists of the loss returns, *i.e.*, the negative log-returns (X_i) on their equity prices at a daily frequency from July 3rd, 2000, to June 30th, 2010. We follow the same set-up as in Cai *et al.* (2015) to extract, for the same time period, daily loss returns (Y_i) of a value-weighted market index aggregating three markets: the New York Stock Exchange, American Express stock exchange and the National Association of Securities Dealers Automated Quotation system.

Cai *et al.* (2015) used $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$, as defined in (27), to estimate the quantile-based MES, $\mathrm{QMES}(\alpha_n) = \mathbb{E}\{X|Y > q_{Y,\alpha_n}\}$, where $\alpha_n = 1 - \frac{1}{n} = 1 - 1/2513$, with two intermediate sequences involved in $\widehat{\gamma}_X$ and $\widehat{\mathrm{QMES}}(\tau_n)$ to be chosen in two steps. Instead, we use our expectile-based method to estimate $\mathrm{QMES}(\alpha_n) \equiv \mathrm{XMES}(\tau_n'(\alpha_n)) = \mathbb{E}\{X|Y > \xi_{Y,\tau_n'(\alpha_n)}\}$, with the same extreme relative frequency α_n that corresponds to a once-per-decade systemic event. We employ the rival estimator $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$ with the same intermediate sequence $\tau_n = 1 - \frac{k}{n}$ in both $\widehat{\gamma}_X$ and $\widehat{\mathrm{QMES}}(\tau_n)$. The conditions required by the procedure were already checked empirically in Cai *et al.* (2015). It only remains to verify that $\gamma_Y < \frac{1}{2}$ as it is the case for γ_X. This assumption is confirmed by the plot of the Hill estimates of γ_Y against the sample fraction k (dashed line) in Figure 8 (a). Indeed, the first stable region appears for $k \in [70, 100]$ with an averaged estimate $\widehat{\gamma}_Y = 0.35$. Hence, by Proposition 5, the estimates $\widehat{\mathrm{XMES}}^{\star}(\alpha_n)$ and $\widetilde{\mathrm{XMES}}^{\star}(\alpha_n)$ are expected to be less extreme than the benchmark values $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$. This is visualised in Figure 18 in the supplement to this article, where the three estimates are graphed as functions of k for each bank. As a matter of fact, both $\widehat{\mathrm{XMES}}^{\star}(\alpha_n)$ and $\widetilde{\mathrm{XMES}}^{\star}(\alpha_n)$ estimate the less extreme risk measure $\mathrm{XMES}(\alpha_n)$ and not the desired intuitive tail measure $\mathrm{XMES}(\tau_n'(\alpha_n)) \equiv \mathrm{QMES}(\alpha_n)$.

The interest here is rather on the composite estimators $\widehat{\mathrm{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$ and $\widetilde{\mathrm{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$, where $\widehat{\mathrm{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$ is actually nothing but $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$. The two rival estimates $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$ and $\widetilde{\mathrm{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$ represent the average daily loss return for a once-per-decade market crisis. They are graphed in Figure 8 (b)-(d) as functions of k for each bank: (b) Goldman Sachs; (c) Morgan Stanley; (d) T. Rowe Price. The first stable region of the plots (b)-(d) appears, respectively, for $k \in [80, 105]$, $k \in [90, 140]$ and $k \in [75, 100]$. The final estimates based on averaging the estimates from these stable regions are reported in the left-hand side of Table 1. It may be seen that both expectile- and quantile-based MES levels for Goldman Sachs and T. Rowe Price are almost equal. However, the MES levels for Morgan Stanley are largely higher than those for Goldman Sachs and T. Rowe Price. It may also be noted that the estimates $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$, obtained here with a single intermediate sequence, are slightly smaller than those obtained in Table 1 of Cai *et al.* (2015) by using two intermediate sequences. Also, these quantile-based estimates appear to be less conservative than our asymmetric least squares-based estimates, but not by much: this minor difference can already be visualised in Figure 8 (b)-(d), where the plots of $\widehat{\mathrm{QMES}}^{\star}(\alpha_n)$, in dashed line, and $\widetilde{\mathrm{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$, in solid line, exhibit a very similar evolution for the three banks.

In our theoretical results we do not enter into the important question of serial dependence.

	Daily loss		Weekly loss	
Bank	$\widetilde{\text{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$	$\widehat{\text{QMES}}^{\star}(\alpha_n)$	$\widetilde{\text{XMES}}^{\star}(\widehat{\tau}_n'(\alpha_n))$	$\widehat{\text{QMES}}^{\star}(\alpha_n)$
Goldman Sachs	0.3123	0.3077	0.3423	0.3375
Morgan Stanley	0.5622	0.5552	0.6495	0.6641
T. Rowe Price	0.3308	0.3098	0.3407	0.3405

Table 1: *Expectile- and quantile-based MES of the three investment banks. The second and third columns report the results based on daily loss returns ($n = 2513$ and $\alpha_n = 1 - 1/n$). The last two columns report the results based on weekly loss returns from the same sample period ($n = 522$ and $\alpha_n = 1 - 1/n$).*

We only consider independent and identically distributed random vectors $(X_1, Y_1), \ldots, (X_n, Y_n)$. One way to reduce substantially the potential serial dependence in this application is by using lower frequency data. As suggested by Cai *et al.* (2015), we choose weekly loss returns in the same sample period. This results in a sample of size $n = 522$. The estimates of γ_Y and $\text{QMES}(\alpha_n) \equiv \text{XMES}(\tau_n'(\alpha_n))$, with $\alpha_n = 1 - \frac{1}{n}$, are displayed in Figure 9 as functions of k. The averaged estimate $\widehat{\gamma}_Y = 0.37$ is obtained from the first stable region $k \in [25, 35]$ of the plot (a). The first stable region of the plots (b)-(d) appears, respectively, for $k \in [27, 36]$, $k \in [23, 33]$ and $k \in [25, 33]$. The final results based on averaging the estimates from these stable regions are reported in the right-hand side of Table 1. They are very similar to those obtained in Cai *et al.* (2015) by resorting to two intermediate sequences. Both expectile- and quantile-based MES estimates are qualitatively robust to the change from daily to weekly data: they are still almost equal for Goldman Sachs and T. Rowe Price, while almost twice higher for Morgan Stanley.

There remains a lot to be done, especially on the extension of our expectile-based methods to a time dynamic setting. Already, Taylor (2008) and Kuan *et al.* (2009) have initiated the use of expectiles to estimate VaR and ES in conditional autoregressive expectile models. The use of expectiles to estimate MES may also work by allowing for dynamics in the covariance matrix via a multivariate GARCH model, similarly to the quantile-based method of Brownlees and Engle (2016). From the perspective of extreme values, one way to deal with the heteroskedasticity present in series of financial returns, similarly to Diebold *et al.* (2000), McNeil and Frey (2000) and McNeil *et al.* (2005, p. 283), is by applying our method to residuals standardized by GARCH conditional volatility estimates. Also, similarly to extreme value analysis under mixing conditions in a univariate setting (see, *e.g.*, Drees (2000)), our extreme value theorems may work under serial dependence with enlarged asymptotic variances.

Acknowledgements

The authors would like to thank three referees, the Associate Editor and the Joint Editor for their valuable suggestions, which have significantly improved the paper. We are also grateful to Juan-Juan Cai and Chen Zhou who kindly sent us their codes about MES numerical illustrations. A. Daouia's research was supported by the Toulouse School of Economics Individual Research Fund (IRF/Daouia-20125).

Supplementary materials

The supplement to this article contains additional simulations, technical lemmas and the proofs of all theoretical results in the main article.

References

[1] Abdous, B. and Remillard, B. (1995). Relating quantiles and expectiles under weighted-symmetry, *Annals of the Institute of Statistical Mathematics*, **47**, 371–384.

[2] Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion, *Journal of Banking & Finance*, **26**, 1505–1518.

[3] Acharya, V.V., Pedersen, L.H., Philippon, T. and Richardson, M. (2012). Measuring systemic risk, *Discussion Paper DP8824*, Centre for Economic Policy Research, London.

[4] Alm, J. (2015). Signs of dependence and heavy tails in non-life insurance data, *Scandinavian Actuarial Journal*, DOI: 10.1080/03461238.2015.1017527.

[5] Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, **9**, 203–228.

[6] Basel Committee (2011). *Basel III: a global regulatory framework for more resilient banks and banking systems.*

[7] Bassett, G.W, Koenker, J.R. and Kordas, G. (2004). Pessimistic Portfolio Allocation and Choquet Expected Utility. *Journal of Financial Econometrics*, **4**, 477–492.

[8] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of extremes: Theory and applications*, Wiley.

[9] Bellini, F. (2012). Isotonicity results for generalized quantiles, *Statistics and Probability Letters*, **82**, 2017–2024.

[10] Bellini, F., Klar, B., Müller, A. and Gianina, E.R. (2014). Generalized quantiles as risk measures, *Insurance: Mathematics and Economics*, **54**, 41–48.

[11] Bellini, F. and Di Bernardino, E. (2015). Risk Management with Expectiles, *The European Journal of Finance*. DOI:10.1080/1351847X.2015.1052150

[12] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.

[13] Breckling, J. and Chambers, R. (1988). M-quantiles, *Biometrika*, **75**, 761–772.

[14] Brownlees, C.T. and Engle, R. (2016). SRISK: A Conditional Capital Shortfall Measure of Systemic Risk, *The Review of Financial Studies*, forthcoming.

[15] Caeiro, F., Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *REVSTAT: Statistical Journal*, **3**, 113–136.

[16] Cai, J., Einmahl, J., de Haan, L. and Zhou, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme, *Journal of the Royal Statistical Society: Series B*, **77**, 417–442.

[17] Chavez-Demoulin, V., Embrechts, P. and Sardy, S. (2014). Extreme-quantile tracking for financial time series, *Journal of Econometrics*, **181**, 44–52.

[18] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer-Verlag, New York.

[19] de Valk, C. (2015). Approximation and estimation of very small probabilities of multivariate extreme events, *Extremes* (to appear).

[20] Diebold, F.X., Schuermann, T. and Stroughair, J.D. (2000). Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management, *Journal of Risk Finance 1* (Winter), 30–36.

[21] Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables, *Ann. Appl. Probab.*, **10**, 1274–1301.

[22] Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function, *Journal of Multivariate Analysis*, **64**, 25–47.

[23] Ehm, W., Gneiting, T., Jordan, A. and Krüger, F. (2016). Of Quantiles and Expectiles: Consistent Scoring Functions, Choquet Representations, and Forecast Rankings, *Journal of the Royal Statistical Society: Series B*, **78**, 505–562.

[24] El Methni, J., Gardes, L. and Girard, S. (2014). Nonparametric estimation of extreme risks from conditional heavy-tailed distributions, *Scandinavian Journal of Statistics*, **41** (4), 988–1012.

[25] El Methni, J. and Stupfler, G. (2016). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica* (to appear).

[26] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer.

[27] Embrechts, P., Puccetti, G. (2007). Aggregating risk across matrix structured loss data: the case of operational risk, *Journal of Operational Risk*, **3**, 29–44.

[28] Engle, R.F., Jondeau, E. and Rockinger, M. (2015). Systemic Risk in Europe, *Review of Finance*, **19**, 145–190.

[29] Fan, J., Gijbels, I. (1996). Local Polynomial Modelling and Its Applications. Monographs on Statistics and Applied Probability, 66. London: Chapman and Hall.

[30] Fissler, T. and Ziegel, J.F. (2016). Higher order elicitability and Osband's principle, *Annals of Statistics*, **4**, 1680–1707.

[31] Geyer, C.J. (1996). On the asymptotics of convex stochastic optimization, unpublished manuscript.

[32] Gilli, M., Këllezi, E. (2006). An Application of Extreme Value Theory for Measuring

Financial Risk, *Computational Economics*, **27**, 207–228.

[33] Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, **106**, 746–762.

[34] Gomes, M.I. and Guillou, A. (2015). Extreme value theory and statistics of univariate extremes: a review, *International Statistical Review*, **83**, 263–292.

[35] Gong, J., Li, Y., Peng, L. and Yao, Q. (2015). Estimation of extreme quantiles for functions of dependent random variables, *Journal of the Royal Statistical Society: Series B*, **77**, 1001–1024.

[36] Hartmann, P., Straetmans, S. and de Vries, C. (2004) Asset market linkages in crisis periods, *Review of Economics and Statistics*, *86*, 313–326.

[37] Hartmann, P., Straetmans, S. and de Vries, C. (2005). Banking system stability. A cross-atlantic perspective, in *The Risks of Financial Institutions, NBER Chapters*, 133–192. National Bureau of Economic Research.

[38] Hartmann, P., Straetmans, S. and de Vries, C. (2010). Heavy tails and currency crises, *Journal of Empirical Finance*, *17*, 241–254.

[39] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1163–1174.

[40] Holzmann, H. and Klar, B. (2016). Expectile asymptotics, *Electronic Journal of Statistics*, **10**, 2355–2371.

[41] Hua, L. and Joe, H. (2011). Second order regular variation and conditional tail expectation of multiple risks, *Insurance: Mathematics and Economics*, **49**, 537–546.

[42] Ibragimov, R., Jaffee, D. and Walden, J. (2009). Nondiversification traps in catastrophe insurance markets, *Review of Financial Studies*, **22**, 959–993.

[43] Knight, K. (1999). Epi-convergence in distribution and stochastic equi-semicontinuity, technical report, University of Toronto.

[44] Koenker, R. and Bassett, G. S. (1978). Regression Quantiles, *Econometrica*, **46**, 33–50.

[45] Krätschmer, V. and Zähle, H. (2016). Statistical inference for expectile-based risk measures, preprint, arXiv:1601.05261.

[46] Kuan, C-M., Yeh, J-H. and Hsu, Y-C. (2009). Assessing value at risk with CARE, the Conditional Autoregressive Expectile models, *Journal of Econometrics*, **2**, 261–270.

[47] Mao, T., Ng, K. and Hu, T. (2015). Asymptotic Expansions of Generalized Quantiles and Expectiles for Extreme Risks, *Probability in the Engineering and Informational Sciences*, **29**, 309–327.

[48] Mao, T. and Yang, F. (2015). Risk concentration based on Expectiles for extreme risks under FGM copula, *Insurance: Mathematics and Economics*, **64**, 429–439.

[49] Martin, R. (2014). Expectiles behave as expected, *Risk Magazine*, **27** (6), 79–83.

[50] McNeil, A.J. and Frey, R. (2000). Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach, *Journal of Empirical Finance*, **7**, 271–300.

[51] McNeil, A.J., Frey, R. and Embrechts, P. (2005). Quantitative Risk Management. New

Jersey: Princeton University Press.

[52] Newey, W.K. and Powell, J.L. (1987). Asymmetric least squares estimation and testing, *Econometrica*, **55**, 819–847.

[53] Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*, Springer.

[54] Rockafellar, R.T. and Uryasev, S. (2002). Conditional Value-at-Risk for General Loss Distributions, *Journal of Banking and Finance*, **26**, 1443–1471.

[55] Schulze Waltrup, L., Sobotka, F., Kneib, T. and Kauermann, G. (2015). Expectile and Quantile Regression - David and Goliath? *Statistical Modelling*, **15**, 433–456.

[56] Sobotka, F. and Kneib, T. (2012). Geoadditive expectile regression, *Computational Statistics and Data Analysis*, **56**, 755–767.

[57] Steenbergen, H.M.G.M., Lassing, B.L., Vrouwenvelder, A.C.W.M. and Waarts, P.H. (2004). Reliability analysis of flood defence systems, *Heron*, **49**, 51–73.

[58] Taylor, J. (2008). Estimating value at risk and expected shortfall using expectiles, *Journal of Financial Econometrics*, **6**, 231–252.

[59] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association*, **73**, 812–815.

[60] Ziegel, J.F. (2016). Coherence and elicitability, *Mathematical Finance*, **26**, 901–918.

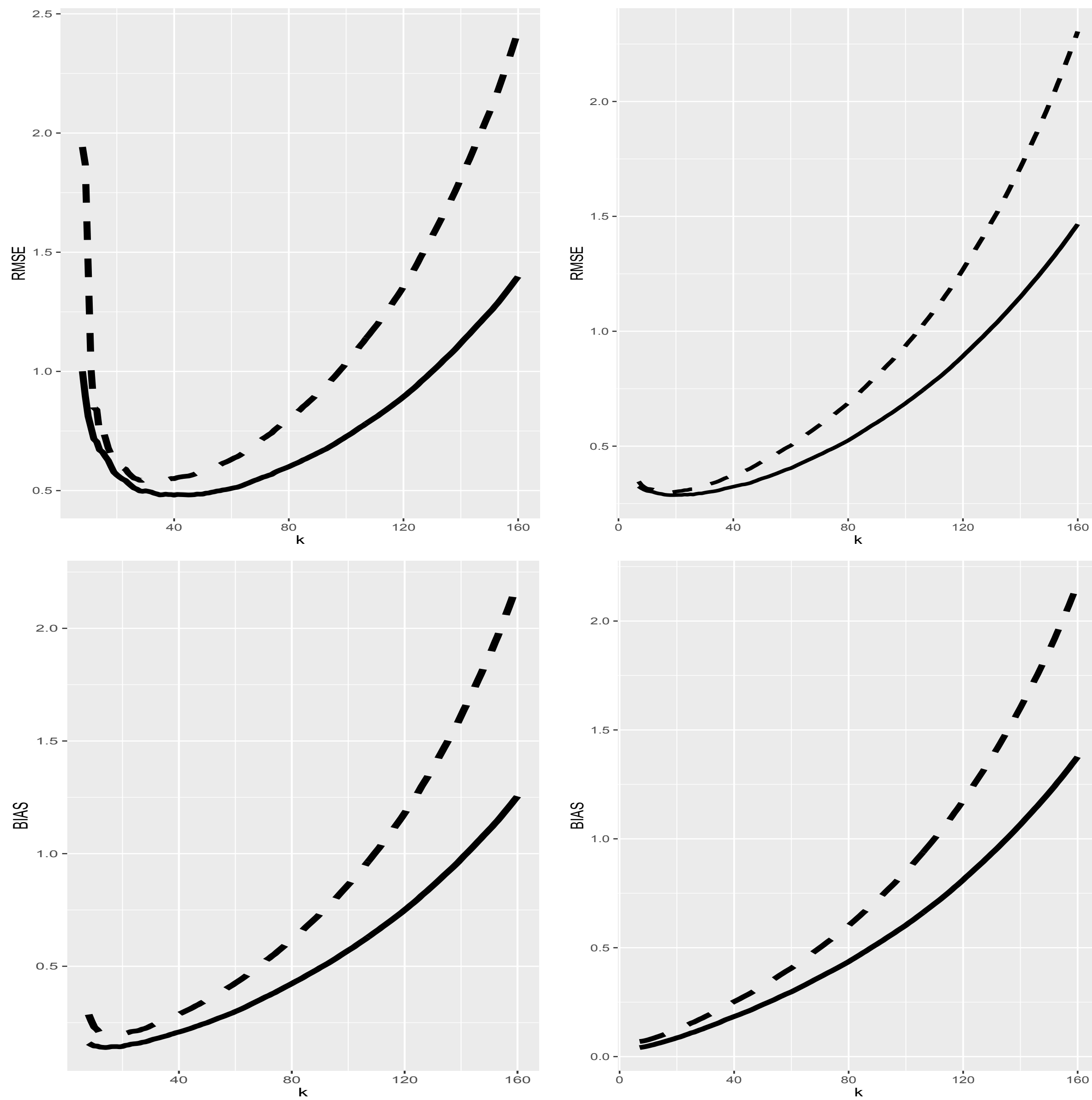

Figure 3: *Root MSE estimates (top panels) and Bias estimates (bottom panels) of* $\widetilde{\xi}^{\star}_{\tau'_n}(k)/\xi_{\tau'_n}$ *(solid line) and* $\widehat{\xi}^{\star}_{\tau'_n}(k)/\xi_{\tau'_n}$ *(dashed line), as functions of* k*, for the* t_3 *and* t_5*-distributions, respectively, from left to right.*

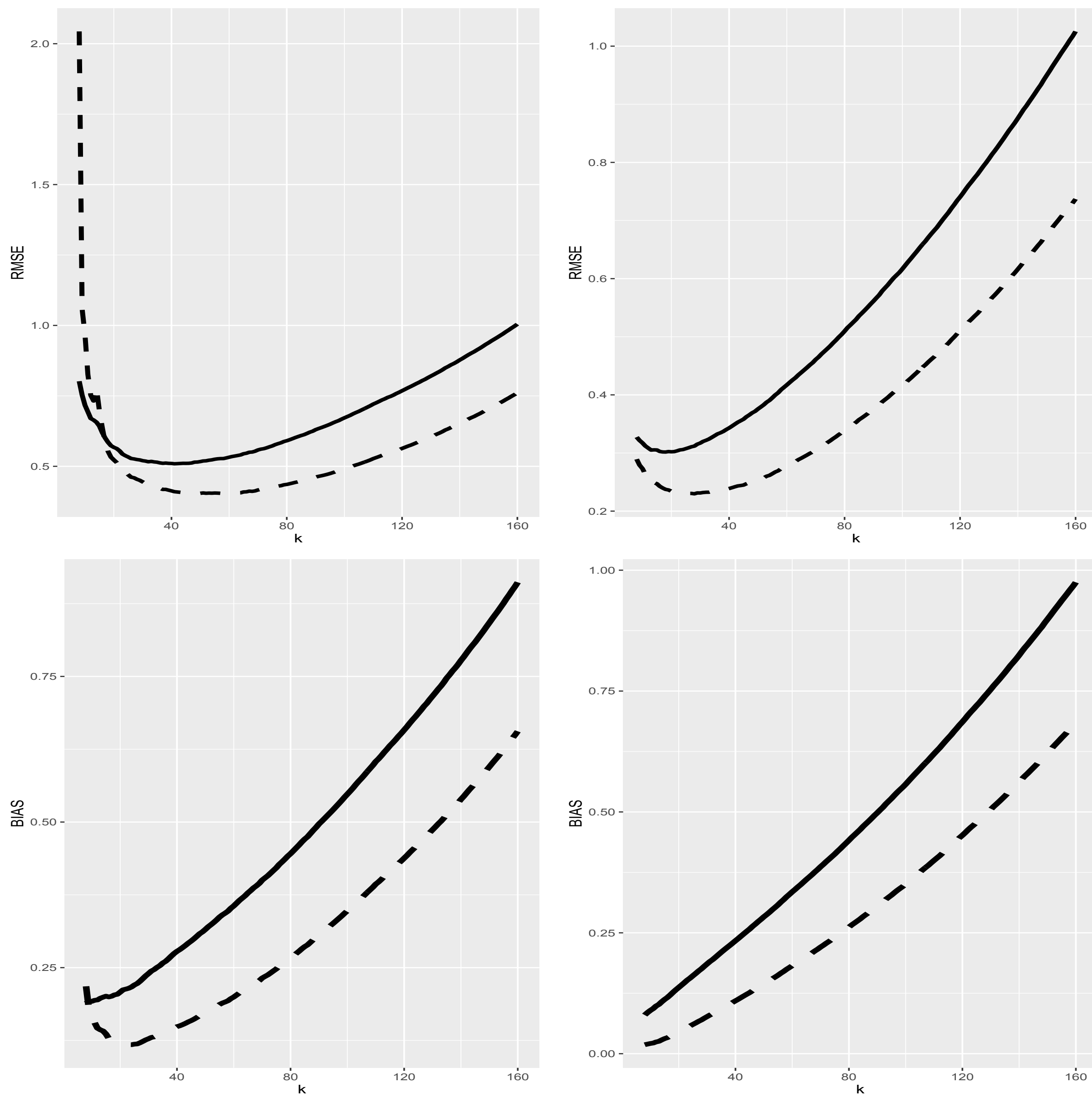

Figure 4: *Root MSE estimates (top panels) and Bias estimates (bottom panels) of* $\widetilde{\xi}^{\star}_{\tau'_n}(k)/\xi_{\tau'_n}$ *(solid line) and* $\widehat{\xi}^{\star}_{\tau'_n}(k)/\xi_{\tau'_n}$ *(dashed line), as functions of* k*, for the positive Student* t_3 *and* t_5*-distributions, respectively, from left to right.*

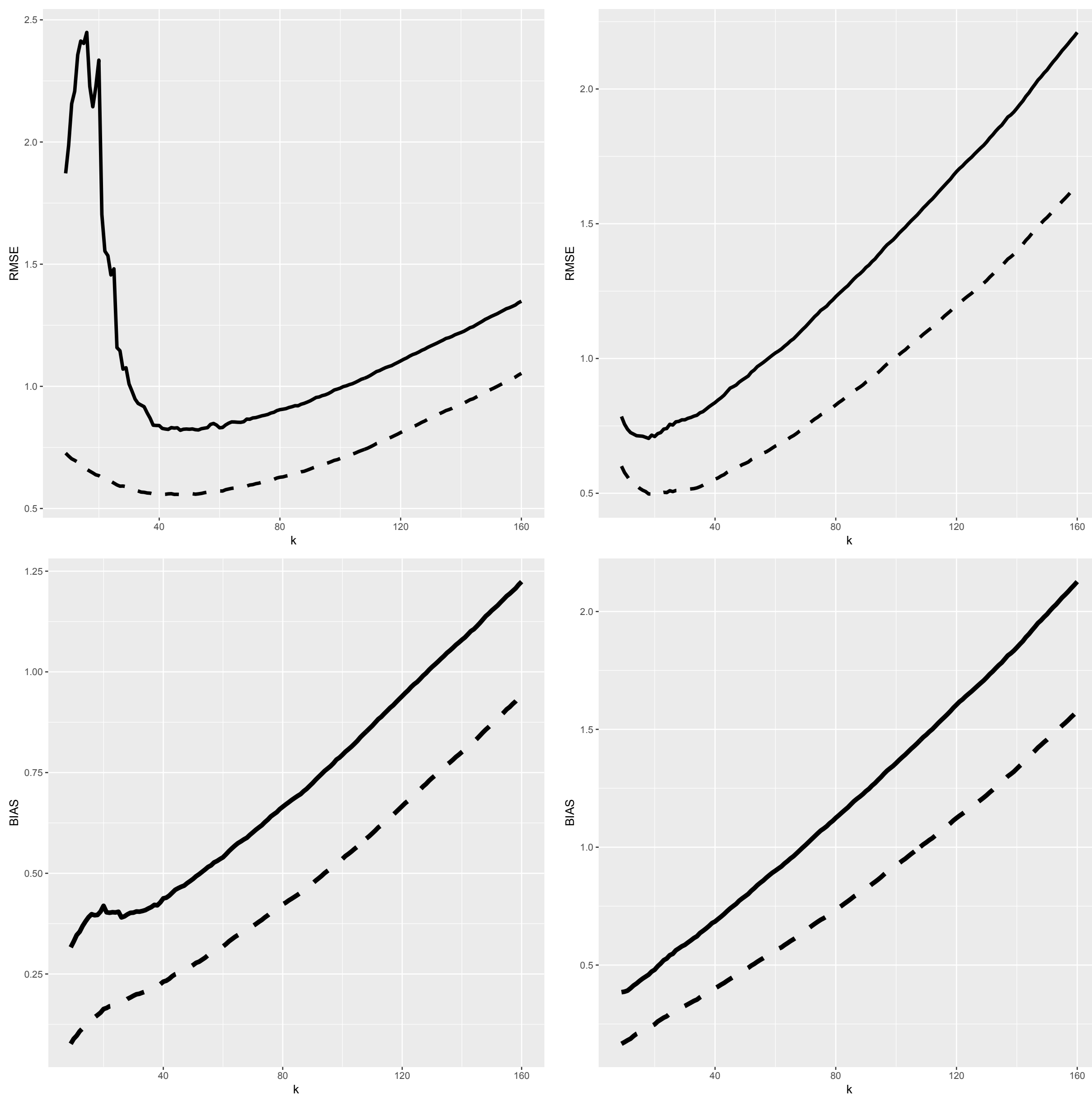

Figure 5: *Root MSE estimates (top panels) and Bias estimates (bottom panels) of* $\widetilde{XMES}^{\star}/XMES$ *(solid line) and* $\widehat{XMES}^{\star}/XMES$ *(dashed line), as functions of* k*, for the bivariate* t_3 *and* t_5*-distributions on* $(0,\infty)^2$*, respectively, from left to right.*

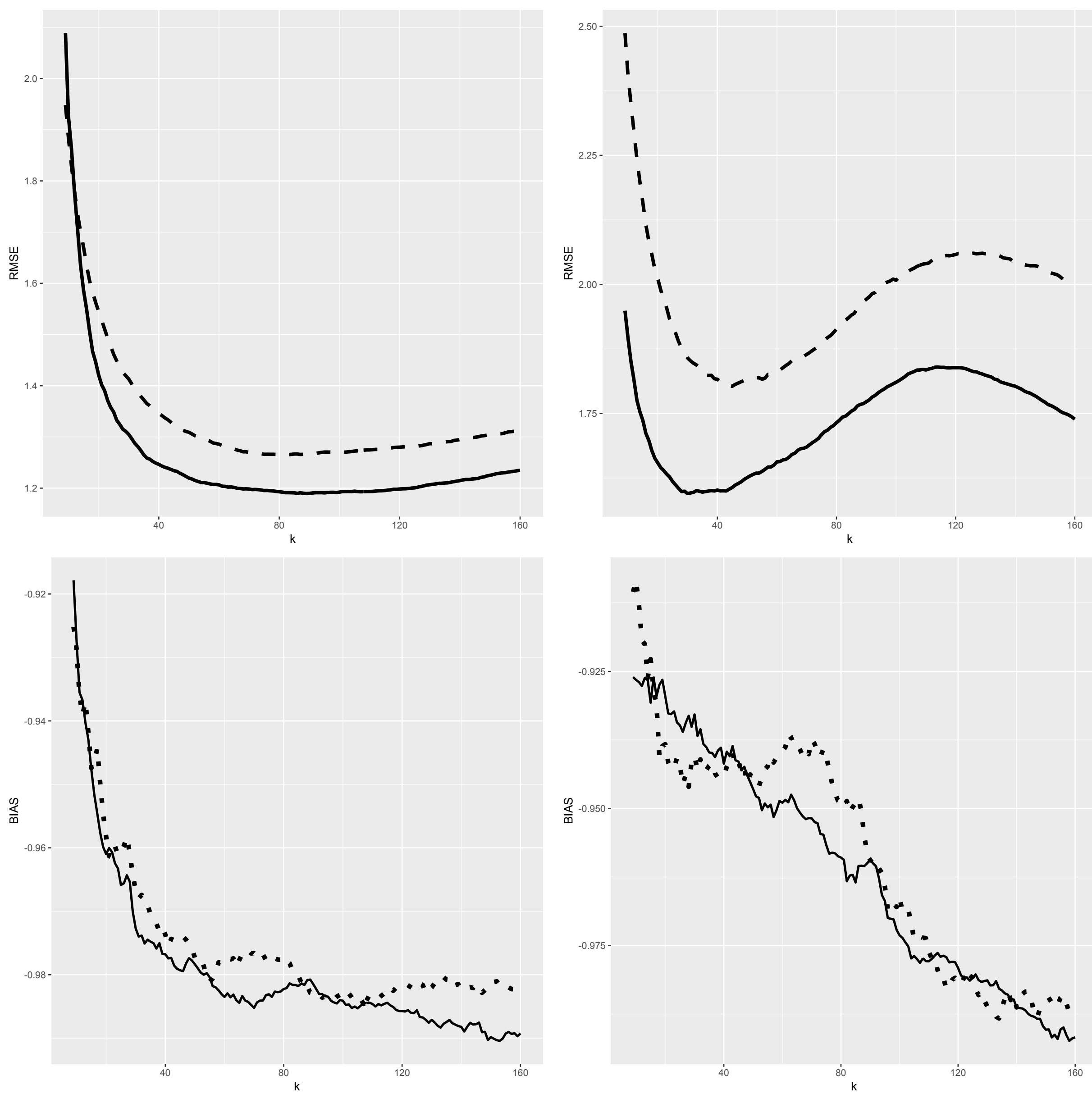

Figure 6: *Root MSE estimates (top panels) and Bias estimates (bottom panels) of* $\widetilde{XMES}^{\star}/XMES$ *(solid line) and* $\widehat{XMES}^{\star}/XMES$ *(dashed line), as functions of* k*, for the transformed* t_3 *and* t_5*-distributions on* $\mathbb{R}^2$*, respectively, from left to right.*

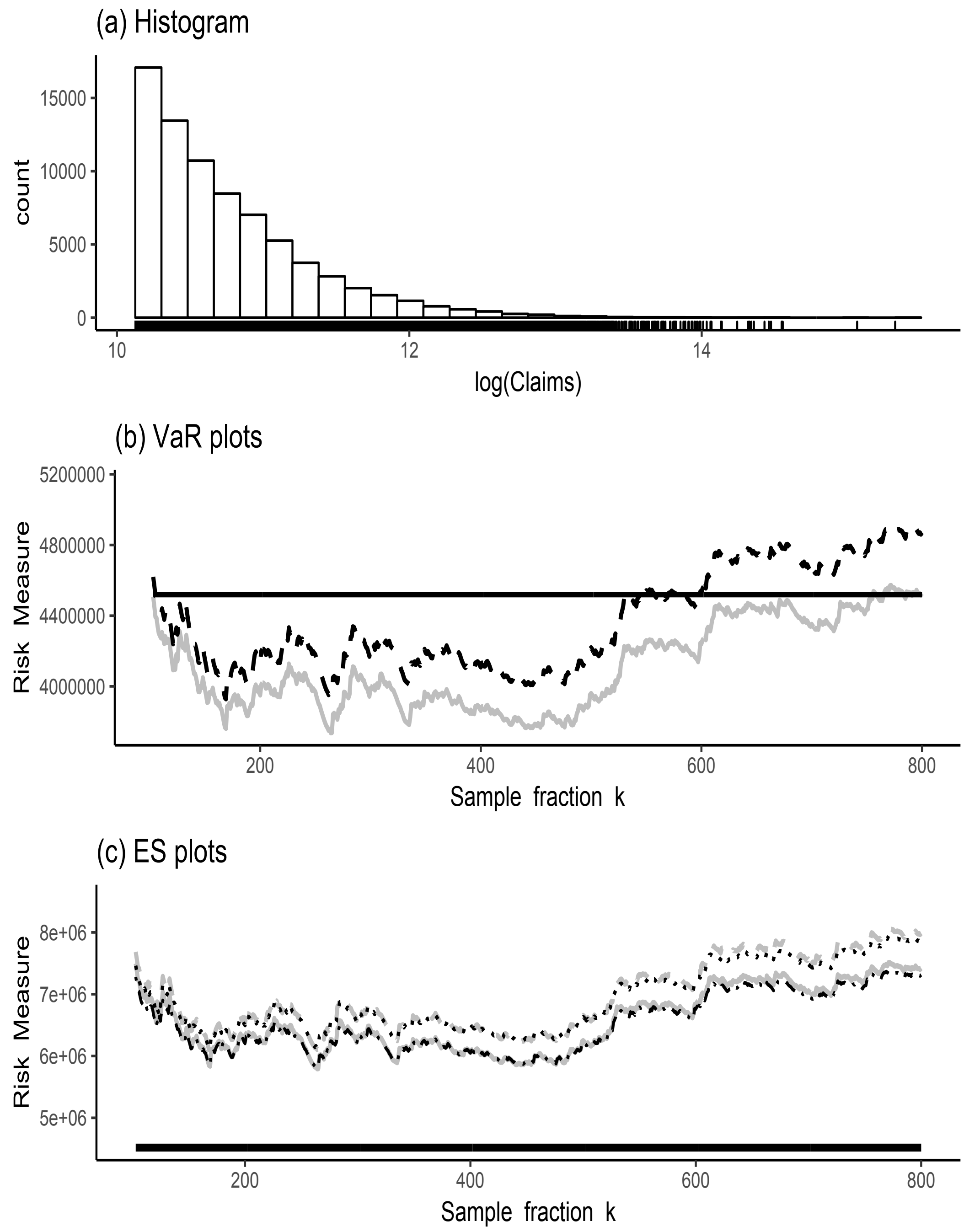

Figure 7: *SOA Group Medical Insurance data. (a) Histogram and scatterplot of the log-claim amounts. (b) The VaR plots* $\{(k,\widetilde{\xi}^{\star}_{\widehat{\tau}'_n(\alpha_n)}(k))\}_k$ *in dashed line and* $\{(k,\widehat{q}^{\star}_{\alpha_n}(k))\}_k$ *in solid line, along with the sample maximum* $Y_{n,n}$ *in horizontal line. (c) The ES plots* $\{(k,\widehat{XES}^{\star}_k(\widehat{\tau}'_n(\alpha_n)))\}_k$ *in solid line,* $\{(k,\widetilde{XES}^{\star}_k(\widehat{\tau}'_n(\alpha_n)))\}_k$ *in dashed line,* $\{(k,\widetilde{XES}^{\dagger}_k(\widehat{\tau}'_n(\alpha_n)))\}_k$ *in dotted line, and* $\{(k,\widehat{QES}^{\star}_k(\alpha_n))\}_k$ *in dashed-dotted, along with* $Y_{n,n}$ *in horizontal line.*

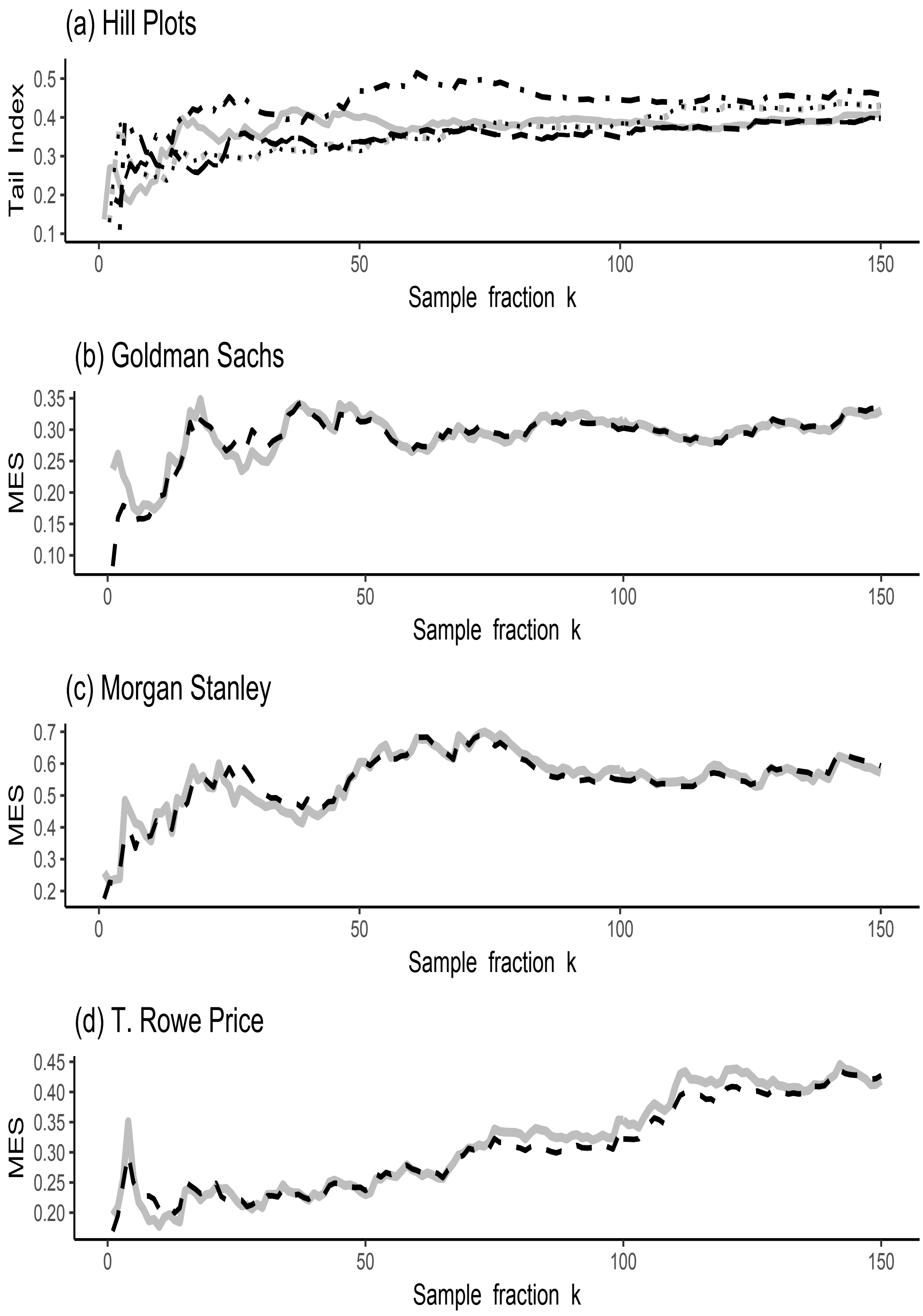

Figure 8: *(a) Hill estimates $\widehat{\gamma}_Y$ based on daily loss returns of market index (dashed), along with $\widehat{\gamma}_X$ based on daily loss returns of three investment banks: Goldman Sachs (solid), Morgan Stanley (dashed-dotted), T. Rowe Price (dotted). (b)-(d) The estimates $\widehat{QMES}^{\star}(\alpha_n)$ in dashed line and $\widetilde{XMES}^{\star}(\widehat{\tau}_n'(\alpha_n))$ in solid line for the three banks, with $n = 2513$ and $\alpha_n = 1 - 1/n$.*

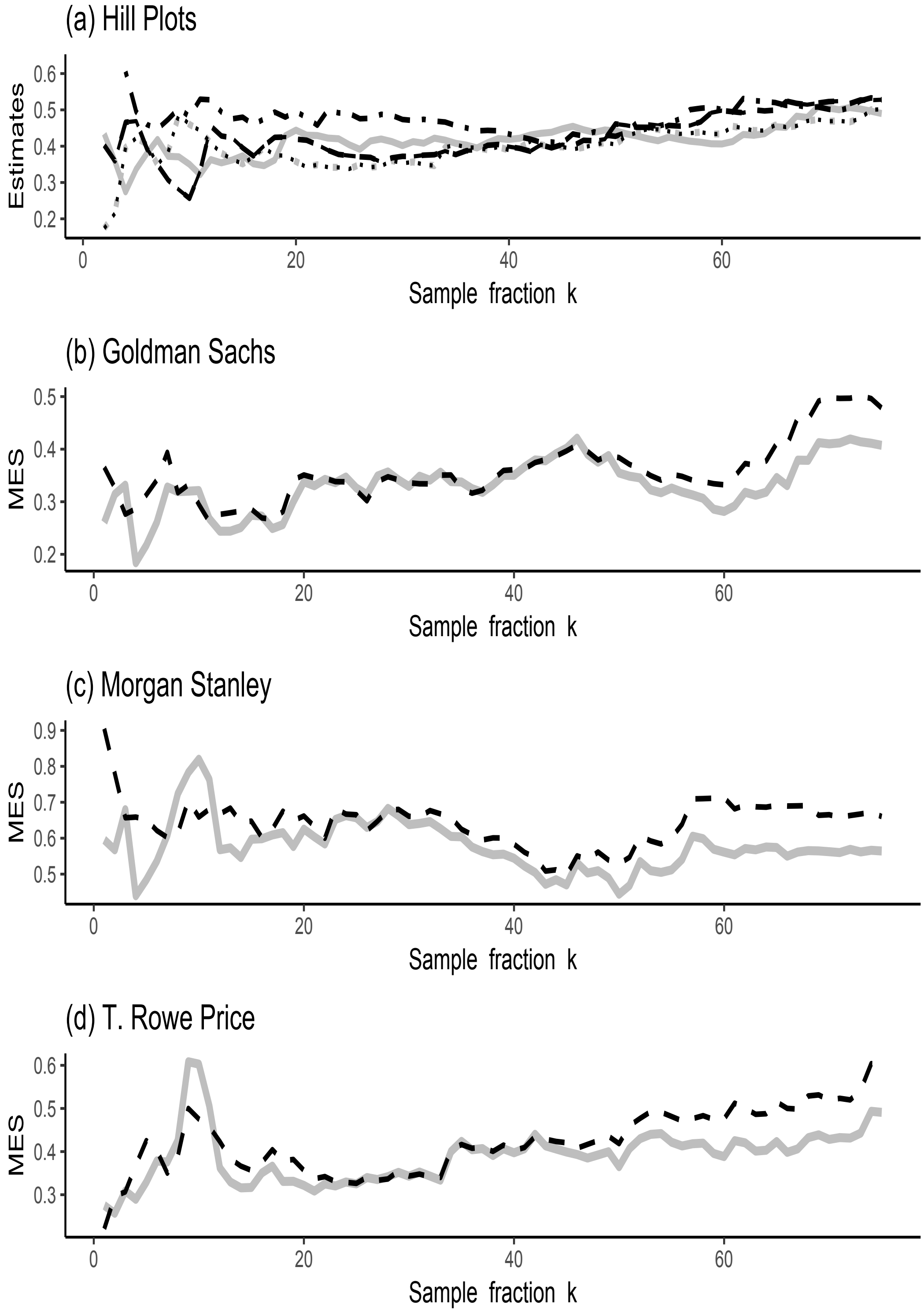

Figure 9: *(a) Hill estimates* $\widehat{\gamma}_Y$ *based on weekly loss returns of market index (dashed), along with* $\widehat{\gamma}_X$ *based on weekly loss returns of three investment banks: Goldman Sachs (solid), Morgan Stanley (dashed-dotted), T. Rowe Price (dotted). (b)-(d) The estimates* $\widehat{QMES}^{\star}(\alpha_n)$ *in dashed line and* $\widetilde{XMES}^{\star}(\widehat{\tau}'_n(\alpha_n))$ *in solid line for the three banks, with* $n = 522$ *and* $\alpha_n = 1 - 1/n$.